



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 184052

TO: Manjunath N Rao
Location: REM-2A01/2C70
Art Unit: 1652
Tuesday, April 11, 2006
Case Serial Number: 09/211691

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518

BOB
barbara.obryen@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

184052
STIC-Biotech/ChemLib

From: Rao, Manjunath N.
Sent: Tuesday, April 04, 2006 8:40 AM
To: STIC-Biotech/ChemLib
Subject: Sequenc esearch request for 09/211691

RECEIVED
APR - 4 2006
STIC/BIOTECH/CHM LIB
(STIC)

From: Manjunath N. Rao
Art Unit 1652, Room 2A01
Mail Box in Room 2C70
Phone: 272-0939

Date: 4-4-06

Please search the following as soon as possible for application with serial number
09/211691

1. SEQ ID NO: 1, 3, 4 against all commercial nucleic acid sequence databases, issued patents/published applications nucleic acid sequence database and pending application nucleic acid sequence database. Please provide a print of results
2. SEQ ID NO: 2 against all commercial amino acid sequence databases, issued patents/published applications amino acid sequence database and pending application amino acid sequence database. Please provide a print of results.

If you have any questions please call me at the above phone number.

Thanks

4/4/2006

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 15:58:52 ; Search time 5660.11 Seconds
(without alignments)
8315.447 Million cell updates/sec

Title: US-09-211-691-1
Perfect score: 828f
Sequence: 1 atgcaaacacagcttcacg.....tcattgtgccttcacataa 828

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:*

1: gb_ba:*

2: gb_in:*

3: gb_env:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pr:*

9: gb_ro:*

10: gb_srb:*

11: gb_sy:*

12: gb_un:*

13: gb_vl:*

14: gb_hg:*

15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	828	100.0	3613	1 NM025839	U25839 Neisseria m
2	826.4	99.8	110000	1 AE002098_20	Continuation (21 o
3	826.4	99.8	349980	6 AX044034	AX044034 Sequence
4	808.8	97.7	4827	1 AY039684	AY039684 Neisseria
5	800.8	96.7	3335	1 AF470659	AF470659 Neisseria
6	791.2	95.6	2959	1 AF240672	AF240672 Neisseria
7	788	95.2	3369	1 AF470660	AF470660 Neisseria
8	786.4	95.0	3367	1 AF470658	AF470658 Neisseria
9	782	94.4	5216	1 AY134877	AY134877 Neisseria
10	775.6	93.7	3289	1 AF470663	AF470663 Neisseria
11	775.6	93.7	3289	1 AF470662	AF470662 Neisseria
12	753.2	91.0	3383	1 AF470655	AF470655 Neisseria
13	750	90.6	3259	1 AF470665	AF470665 Neisseria
14	748.4	90.4	3066	1 AY134878	AY134878 Neisseria
15	746.8	90.2	349061	1 NMA222491	AL162753 Neisseria
16	740	89.4	3266	1 AF470661	AF470661 Neisseria
17	736.8	89.0	3136	1 AF470656	AF470656 Neisseria
18	736.8	89.0	3287	1 AF470657	AF470657 Neisseria

19	735.2	88.8	4517	1 AF355193	AF355193 Neisseria
20	726	87.7	5859	1 NGU14554	U14554 Neisseria g
21	726	87.7	5859	6 AR025350	AR025350 Sequence
22	726	87.7	5859	6 AR025351	AR025351 Sequence
23	726	87.7	5859	6 AR111460	AR111460 Sequence
24	726	87.7	5859	6 I24768	I24768 Sequence 1
25	726	87.7	5859	6 I24769	I24769 Sequence 7
26	726	87.7	5859	6 AR183924	AR183924 Sequence
27	726	87.7	5859	6 AR183925	AR183925 Sequence
28	726	87.7	5859	6 AR207874	AR207874 Sequence
29	726	87.7	5859	6 AR567686	AR567686 Sequence
30	726	87.7	5859	6 AR567687	AR567687 Sequence
31	721.2	87.1	53922	1 AB004969_21	Continuation (22 o
32	710	85.7	4883	1 AF313394	AF313394 Neisseria
33	708.4	85.6	5262	1 AY134876	AY134876 Neisseria
34	706	85.3	4862	1 AF121135	AF121135 Neisseria
35	600.2	72.5	3851	1 AY039686	AY039686 Neisseria
36	549	66.3	3851	1 AY039685	AY039685 Neisseria
37	531.4	64.2	2343	1 AF355194	AF355194 Neisseria
38	523.4	63.2	2243	1 AF470664	AF470664 Neisseria
39	484.6	58.5	843	1 AF208063	AF208063 Neisseria
40	459.6	55.5	3712	1 NM065788	U65788 Neisseria m
41	426	51.4	840	6 AX565941	AX565941 Sequence
42	115.6	14.0	10827	1 AE006155	AE006155 Pasteurel
43	103.2	12.5	3998	1 AF224466	AF224466 Haemophil
44	103.2	12.5	301191	1 AE017152	AE017152 Haemophil
45	94.2	11.4	807	1 AY091470	AY091470 Haemophil

ALIGNMENTS

RESULT 1	NM025839	3613 bp	DNA	linear	BCT 18-MAY-1996
LOCUS	NM025839				
DEFINITION	Neisseria meningitidis glycy1 tRNA synthetase (glys) gene, partial cds, and glycosyl transferase (lgta), (lgta), (lgta) genes,				
ACCESSION	U25839				
VERSION	U25839.1	GI:973183			
KEYWORDS					
SOURCE	U25839				
ORGANISM	Neisseria meningitidis				
REFERENCE	Neisseria meningitidis				
AUTHORS	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.				
TITLE	1 (bases 1 to 3613)				
JOURNAL	Jennings,M.P., Hood,D.W., Peak,I.R., Virji,M. and Moxon,R.R.				
PUBMED	Molecular analysis of a locus for the biosynthesis and phase-variable expression of the lacto-N-neotetraose terminal lipopolysaccharide structure in Neisseria meningitidis Mol. Microbiol. 18 (4), 729-740 (1995)				
REFERENCE	8817494				
AUTHORS	2 (bases 1 to 3613)				
TITLE	Jennings,M.P.				
JOURNAL	Direct Submission				
FEATURES	Submitted (27-APR-1995) Michael P. Jennings, Department of Pediatrics, University of Oxford, Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford, Oxfordshire OX3, 9DU, United Kingdom				
source	location/Qualifiers				
gene	1..3613				
CDS	/organism="Neisseria meningitidis"				
	/mol_type="genomic DNA"				
	/strain="MC58"				
	/db_xref="taxon:487"				
	1..264				
	/gene="glys"				
	<1..264				
	/gene="glys"				
	/codon_start=1				
	/transl_table=11				
	/product="glycy1 tRNA synthetase"				
	/protein_id="AAC44083.1"				

```
/db_xref="GI:973184"
/translation="SLAQDEEKALFAAAGQIQPIAAVAAGNFQALSELASVQPO
VAFPPGVVWMEADAIVKONRLNLRLEQNNAAVADIALGE"
/gene=".1329
/gene="1gta"
318..321
/gene="1gta"
328..329
/gene="1gta"
/function="involved in the biosynthesis of the terminal
lacto-N-neotetraose LPS structure"
/codon_start=1
/evidence=experimental
/translation=11
/product="glycosyl transferase"
/protein_id="AAC44084.1"
/db_xref="GI:973185"
/translation="MQLVSVLICAVNVEKPAOSILAAVNTWENLILIVDSDTD
GLTALAEQEDODGRLILAEPRNSGLIPSLNIGDLKAGGGGAYIARTDIDIA
PMIETKIVGEMKRSIIAMGAMLEVLSEKDNRLARHBEKTKWKPTRREDIAD
PFGNPIHNTMTMRSVTDGRLRTNTEADNEDYQFWDVSKGLATYPAIVKIR
LHANGVSKTISIRHSLAQGIQKTANDPLQSGMFKTRPDSLEYQIAVAVALLEKH
LPEDFELARRRLYOCFKRTDTLPAGAWLDFADGRMRLFTLRQYFGLHRLKNR"
582..595
/gene="1gta"
/notes="encodes homopolymeric tract of 14 guanosine
residues which control the phase variable expression of
the 1gta gene via a slipped-strand mispairing mechanism"
1361..2198
/gene="1gta"
1361..1364
/gene="1gta"
1371..2198
/gene="1gta"
/function="involved in the biosynthesis of the terminal
lacto-N-neotetraose LPS structure"
/codon_start=1
/evidence=experimental
/translation=11
/product="glycosyl transferase"
/protein_id="AAC44085.1"
/db_xref="GI:973186"
/translation="MGNHVISLASAERRAHIAITFGNHLIPQFPDMLMSEELBOA
MAELVGLAKQHLSEVEKACPMASHAVLMKQALDGLPYITFEDVILGSEEFELA
EDAMLOBERDPDTAFIVLEETMFAKIVRBDKLVNTENSFLBSEHGTAGYIISREAMR
FLERPAVLPAIVDAVDMVMFTYVFDKGMVYOVNVALCTQELHAKFLSKNMLG
SDLEKROERHRSRLKVMFLKRALGFGREKKKMRBQRAELKAYGSRVTSFK"
2337..3177
/gene="1gta"
2337..2340
/gene="1gta"
2347..3177
/gene="1gta"
/function="involved in the biosynthesis of the terminal
lacto-N-neotetraose LPS structure"
/codon_start=1
/evidence=experimental
/translation=11
/product="glycosyl transferase"
/protein_id="AAC44086.1"
/db_xref="GI:973187"
/translation="MGNHVISLASAERRAHIAITFGNHLIPQFPDMLMSEELNR
MAELVGLAKQHLSEVEKACPMASHAVLMKQALDGLPYITFEDVILGSEEFELA
EDAMLOBERDPDTAFIVLEETMFAKIVRBDKLVNTENSFLBSEHGTAGYIISREAMR
FLERPAVLPAIVDAVDMVMFTYVFDKGMVYOVNVALCTQELHAKFLSKNMLG
SDLEKROERHRSRLKVMFLKRALGFGREKKKMRBQRAELKAYGSRVTSFK"
3297..3369
repeat_region
repeat_region
3538..3606
ORIGIN
Query Match 100.0%; Score 828; DB 1; Length 3613;
Best Local Similarity 100.0%; Pred. No. 9.1e-187;
```

```
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCAAAACCAAGTATGAGTTAGCTTCCGCCGAGAAACGAGGCGCAATTCGCCAT 60
Db 1371 ATGCAAAACCAAGTATGAGTTAGCTTCCGCCGAGAAACGAGGCGCAATTCGCCAT 1430
QY 61 ACCTTGGGAGGACGAGCATCCGTTTCAGATTTTTCAGACGACATGATCCGCTGAANAAG 120
Db 1431 ACCTTGGGAGGACGAGCATCCGTTTCAGATTTTTCAGACGACATGATCCGCTGAANAAG 1490
QY 121 CTGGAACGAGCAATGCGGAACTCGTCCCGGCTTGTGGGCAACCCCTATTGAGCGGA 180
Db 1491 CTGGAACGAGCAATGCGGAACTCGTCCCGGCTTGTGGGCAACCCCTATTGAGCGGA 1550
QY 181 GTGAAAAAGCCCTGTTATGAGCCAGCCGATTTGTGAAAGCAGCAATTGACGAAAGT 240
Db 1551 GTGAAAAAGCCCTGTTATGAGCCAGCCGATTTGTGAAAGCAGCAATTGACGAAAGT 1610
QY 241 CTGCGGTATATCAACGTTATTTGAGAGCAGTTTTTACTGGCGAAGGTGAGAAAAATTC 300
Db 1611 CTGCGGTATATCAACGTTATTTGAGAGCAGTTTTTACTGGCGAAGGTGAGAAAAATTC 1670
QY 301 CTTGCCGAAGACGCTTGCTGCAAGAACGCTTTGACCCGATACCGCTTTATCGTCCGC 360
Db 1671 CTTGCCGAAGACGCTTGCTGCAAGAACGCTTTGACCCGATACCGCTTTATCGTCCGC 1730
QY 361 TTGAAAACGATGTTTATGACAGTCTGACCTTCCGCTTGGCGGATTACTGCGGG 420
Db 1731 TTGAAAACGATGTTTATGACAGTCTGACCTTCCGCTTGGCGGATTACTGCGGG 1790
QY 421 CGCGCTTTCGCTGTTGAAAGCAACACATGAGGAGGAGCGAGGCTATATCATTTCCGA 480
Db 1791 CGCGCTTTCGCTGTTGAAAGCAACACATGAGGAGGAGCGAGGCTATATCATTTCCGA 1850
QY 481 AAAGCATGCGGTTTCTTCTGACAGGTTTCCGCGCTGCGCCGAAAGGCTGACCC 540
Db 1851 AAAGCATGCGGTTTCTTCTGACAGGTTTCCGCGCTGCGCCGAAAGGCTGACCC 1910
QY 541 GTGATCTGATGATGTTGAGGATTTTTCAGACAGGAAAGAAATGCGGTTTGCAGCTC 600
Db 1911 GTGATCTGATGATGTTGAGGATTTTTCAGACAGGAAAGAAATGCGGTTTGCAGCTC 1970
QY 601 AATCCCGCTTGTGGGCCCAAGAGCTGATATGCAAGTTTTCAGACCAAAACAGCGCA 660
Db 1971 AATCCCGCTTGTGGGCCCAAGAGCTGATATGCAAGTTTTCAGACCAAAACAGCGCA 2030
QY 661 TTGGGCAAGCTGATGAAACAGACCGGCTCTGAAACCGCAACAGCAAAAGCGGATTCC 720
Db 2031 TTGGGCAAGCTGATGAAACAGACCGGCTCTGAAACCGCAACAGCAAAAGCGGATTCC 2090
QY 721 CCCGCAACATTCAAACACCGCTGATCCGCGCTTGAACCAATATCAGAGGAAAGG 780
Db 2091 CCCGCAACATTCAAACACCGCTGATCCGCGCTTGAACCAATATCAGAGGAAAGG 2150
QY 781 GAAAAACGCGCGCAAAAGCGCAACAGTTCATTGCTTTCATATA 828
Db 2151 GAAAAACGCGCGCAAAAGCGCAACAGTTCATTGCTTTCATATA 2198
RESULT 2
AB002098_20/c
Sequence split into 23 fragments LOCUS AB002098 Accession AB002098
Fragment Name Begin End
AB002098_00 1 11000
AB002098_01 100001 210000
AB002098_02 200001 310000
AB002098_03 300001 410000
AB002098_04 400001 510000
AB002098_05 500001 610000
AB002098_06 600001 710000
AB002098_07 700001 810000
AB002098_08 800001 910000
```

AE002098_09	900001	1010000			
AE002098_10	1000001	1110000			
AE002098_11	1100001	1210000			
AE002098_12	1200001	1310000			
AE002098_13	1300001	1410000			
AE002098_14	1400001	1510000			
AE002098_15	1500001	1610000			
AE002098_16	1600001	1710000			
AE002098_17	1700001	1810000			
AE002098_18	1800001	1910000			
AE002098_19	1900001	2010000			
AE002098_20	2000001	2110000			
AE002098_21	2100001	2210000			
AE002098_22	2200001	2272360			
Continuation (21 of 23) of AE002098 from base 2000001 (AE002098 Neisseria meningitidis W					
Query Match	99.8%;	Score 826.4;	DB 1;	Length 110000;	
Best Local Similarity	99.9%;	Pred. No. 2.5e-186;			
Matches 827;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
QY	1	ATGCAAAACACGTTATCAGTTAGCTTCCGCGCAGAAAGCAGGCGGCAATTGCCGAT	60		
DB	26697	ATGCAAAACACGTTATCAGTTAGCTTCCGCGCAGAAAGCAGGCGGCAATTGCCGAT	26638		
QY	61	ACCTTCGCGACGACGCGCATCCGTTCACTTTTTCAGCTGATGCGCTGAAAGG	120		
DB	26637	ACCTTCGCGACGACGCGCATCCGTTCACTTTTTCAGCTGATGCGCTGAAAGG	26578		
QY	121	CTGGAACAGGCAATGCGGAATCTGTCGCCGCTTGTCGCGCACTTATTTGAGCGGA	180		
DB	26577	CTGGAACAGGCAATGCGGAATCTGTCGCCGCTTGTCGCGCACTTATTTGAGCGGA	26518		
QY	181	GTCGAAAAGCCGCTTTATGAGCAGCGCTATGTCGAAAGCAGGATGTCGAAAGT	240		
DB	26517	GTCGAAAAGCCGCTTTATGAGCAGCGCTATGTCGAAAGCAGGATGTCGAAAGT	26458		
QY	241	CTGCCGTATATCAACGTAATTTGAGACGAGTTTACTCGCGCAAGGTGAGAAAAATTC	300		
DB	26457	CTGCCGTATATCAACGTAATTTGAGACGAGTTTACTCGCGCAAGGTGAGAAAAATTC	26398		
QY	301	CTTGCAGAGACGCTTGCTGCAAGAAAGCTTTGACCCGGAATACCGCTTTATCTGTCGC	360		
DB	26397	CTTGCAGAGACGCTTGCTGCAAGAAAGCTTTGACCCGGAATACCGCTTTATCTGTCGC	26338		
QY	361	TTGAAAACGATGTTATGACAGCTCTGACCTCCGCGCTGCGGAGATTAATCTCGCGG	420		
DB	26337	TTGAAAACGATGTTATGACAGCTCTGACCTCCGCGCTGCGGAGATTAATCTCGCGG	26278		
QY	421	CGGCGCTTTCGCTGTTGAAAGAGAACTGCGGAGGCTATATCATTTCCCGA	480		
DB	26277	CGGCGCTTTCGCTGTTGAAAGAGAACTGCGGAGGCTATATCATTTCCCGA	26218		
QY	481	AAAGCGATGCGGTTTCTCTGACAGGTTTCCGCGCTGCGCGCGAAGGGCTGCAATCC	540		
DB	26217	AAAGCGATGCGGTTTCTCTGACAGGTTTCCGCGCTGCGCGCGAAGGGCTGCAATCC	26158		
QY	541	GTCGATCTGATGATGTTTCAAGCAATTTTTCAGACAGGAGAAATGCGGTTTCCAGCTC	600		
DB	26157	GTCGATCTGATGATGTTTCAAGCAATTTTTCAGACAGGAGAAATGCGGTTTCCAGCTC	26098		
QY	601	AATCCGCGCTTGCGCGCAAGAGCTGATTTATGCAAGTTTTCAGCAAGCAAGGCGCA	660		
DB	26097	AATCCGCGCTTGCGCGCAAGAGCTGATTTATGCAAGTTTTCAGCAAGCAAGGCGCA	26038		
QY	661	TTGCGGACGCTGATCGAAGCAAGCCGCTCTGAAACGCAAGCAAGGCGCGATTC	720		
DB	26037	TTGCGGACGCTGATCGAAGCAAGCCGCTCTGAAACGCAAGCAAGGCGCGATTC	25978		
QY	721	CCGCGCAACATTTCAACACCGCTGATCCGCGCTTTCAGCAAAATCAGACAGGAAAG	780		
DB	25977	CCGCGCAACATTTCAACACCGCTGATCCGCGCTTTCAGCAAAATCAGACAGGAAAG	25918		
QY	781	GAAGAAAGCGCGCAAGAGCGCAAGCTTATGTCGCTTTCAGATTA	828		

Db	25917	GAAGAAAGCGCGCAAGAGCGCGCAAGCTTATGTCGCTTTCAGATTA	25870		
RESULT 3					
AX044034/c		349980 bp	DNA	linear	PAT 24-NOV-2000
LOCUS		Sequence 113 from Patent WO0066791.			
DEFINITION					
ACCESSION	AX044034				
VERSION	AX044034.1	GI:11342918			
KEYWORDS					
SOURCE		Neisseria meningitidis			
ORGANISM		Neisseria meningitidis			
REFERENCE		Bacteria: Proteobacteria; Betaproteobacteria; Neisseriales;			
AUTHORS		Neisseriaceae; Neisseria.			
TITLE					
JOURNAL					
FEATURES					
source		1. 349980			
		/organism="Neisseria meningitidis"			
		/mol_type="unassigned DNA"			
		/db_xref="taxon:487"			
		/note="sequence too long, cut in 8 pieces."			
		seq 1: 1 to 349980 349980 bases			
		seq 108: 300001 to 649980 349980 bases			
		seq 109: 600001 to 949980 349980 bases			
		seq 110: 900001 to 1249980 349980 bases			
		seq 111: 1200001 to 1549980 349980 bases			
		seq 112: 1500001 to 1849980 349980 bases			
		seq 113: 1800001 to 2149980 349980 bases			
		seq 114: 2100001 to 2272325 172325 bases"			
ORIGIN					
Query Match	99.8%;	Score 826.4;	DB 6;	Length 349980;	
Best Local Similarity	99.9%;	Pred. No. 2.6e-186;			
Matches 827;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
QY	1	ATGCAAAACACGTTATCAGTTAGCTTCCGCGCAGAAAGCAGGCGGCAATTGCCGAT	60		
DB	22656	ATGCAAAACACGTTATCAGTTAGCTTCCGCGCAGAAAGCAGGCGGCAATTGCCGAT	226597		
QY	61	ACCTTCGCGACGACGCGCATCCGTTCACTTTTTCAGCTGATGCGCTGAAAGG	120		
DB	22656	ACCTTCGCGACGACGCGCATCCGTTCACTTTTTCAGCTGATGCGCTGAAAGG	226537		
QY	121	CTGGAACAGGCAATGCGGAATCTGTCGCCGCTTGTCGCGCACTTATTTGAGCGGA	180		
DB	22656	CTGGAACAGGCAATGCGGAATCTGTCGCCGCTTGTCGCGCACTTATTTGAGCGGA	226477		
QY	181	GTCGAAAAGCCGCTTTATGAGCAGCGCTATGTCGAAAGCAGGATGTCGAAAGT	240		
DB	22647	GTCGAAAAGCCGCTTTATGAGCAGCGCTATGTCGAAAGCAGGATGTCGAAAGT	226417		
QY	241	CTGCCGTATATCAACGTAATTTGAGACGAGTTTACTCGCGCAAGGTGAGAAAAATTC	300		
DB	22641	CTGCCGTATATCAACGTAATTTGAGACGAGTTTACTCGCGCAAGGTGAGAAAAATTC	226357		
QY	301	CTTGCAGAGACGCTTGCTGCAAGAAAGCTTTGACCCGGAATACCGCTTTATCTGTCGC	360		
DB	22635	CTTGCAGAGACGCTTGCTGCAAGAAAGCTTTGACCCGGAATACCGCTTTATCTGTCGC	226297		
QY	361	TTGAAAACGATGTTATGACAGCTCTGACCTCCGCGCTTTCAGCAAAATCAGACAGG	420		
DB	22629	TTGAAAACGATGTTATGACAGCTCTGACCTCCGCGCTTTCAGCAAAATCAGACAGG	226237		
QY	421	CGGCGCTTTCGCTGTTGAAAGAGAACTGCGGAGGCTATATCATTTCCCGA	480		
DB	22623	CGGCGCTTTCGCTGTTGAAAGAGAACTGCGGAGGCTATATCATTTCCCGA	226177		

481 AAAGCGATGCGGTTTCTTCTGACAGGTTTGGCCGCTGCGCGCGAAGGCGTGCACCCC 540
Db 226176 AAAGCGATGCGGTTTCTTCTGACAGGTTTGGCCGCTGCGCGCGAAGGCGTGCACCCC 226117
Qy 541 GTCGATCTGATGATGTTTTCAGCGATTTTTCGACAGGGAAGAAATGCCGTTTTCAGAGCTC 600
Db 226116 GTCGATCTGATGATGTTTTCAGCGATTTTTCGACAGGGAAGAAATGCCGTTTTCAGAGCTC 226057
Qy 601 AATCCCGGCTTGGCCGCGCGAAGGCGTGCATTAAGCCAACTTTTTCAGACCCAAACAGCGCA 660
Db 226056 AATCCCGGCTTGGCCGCGCGAAGGCGTGCATTAAGCCAACTTTTTCAGACCCAAACAGCGCA 225997
Qy 661 TTGGGAGCGCTGATTCGACACGACCGCTCTTGAACCGCAACAGCAAGGCGCGATTC 720
Db 225996 TTGGGAGCGCTGATTCGACACGACCGCTCTTGAACCGCAACAGCAAGGCGCGATTC 225937
Qy 721 CCCGCCAATTCGACATTCGACACCGCTGATTCGCGCTTGAACCAAAATTCAGAGGAAAG 780
Db 225936 CCCGCCAATTCGACATTCGACACCGCTGATTCGCGCTTGAACCAAAATTCAGAGGAAAG 225877
Qy 781 GAAAAACGCGCGCAAGGCGCGAAGGCGTGCATTCGCGCTTTCGCAATTA 828
Db 225876 GAAAAACGCGCGCAAGGCGCGAAGGCGTGCATTCGCGCTTTCGCAATTA 225829

RESULT 4 4827 bp DNA linear BCT 02-JAN-2003
LOCUS AY039684
DEFINITION Neisseria meningitidis strain B305-Tr4 lipooligosaccharide
ACCESSION AY039684
VERSION AY039684.1 GI:21654774
KEYWORDS
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 4827)
AUTHORS Zhu, P., Tsai, C.M. and Frasch, C.E.
JOURNAL Immunologic and genetic characterization of lipooligosaccharide
PUBMED variants in a Neisseria meningitidis serogroup C strain
12423771
TITLE Direct Submision
AUTHORS Zhu, P., Lacerda, M.D., Tsai, C.M. and Frasch, C.E.
JOURNAL Submitted (12-JUN-2001) LBP/DBPAP, CBER/FDA, 8800 Rockville Pike,
Bethesda, MD 20892, USA
FEATURES
source Location/Qualifiers
1..4827
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="B305-Tr4"
/db_xref="taxon:487"
38..439
/note="OrfA; allele 2"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK85136.1"
/db_xref="GI:21654775"
/translation="MPESEAFRRHAYRENKLOPLSVLICAVNVEKYFAQSLAAVYNO
TWRNLEILVDDGSTGDTALAKDROKDSRKLIAHLRLNKKQQRDPSANTPKHR
LIRALTKISREKRRQRREGLIKIIVPQ"
451..1383
/gene="1gta"
/gene="1gta"
/allele="6"
451..1383
/gene="1gta"
/allele="6"
/note="1gta"

/codon_start=1
/transl_table=11
/product="alpha 1,4 galactosyltransferase"
/protein_id="AAK85137.1"
/db_xref="GI:21654776"
/translation="MDIVFAADNDVAAVLCVAAKSVAAHPDTEIRPHVLDAI;SEAN
RAAVANLRGGGNTIRFDVNPEDPAGPPLNIRH;SITTYARLKI;GEYADDDKYLIL
IDVLRDLSKRLMTDLDNMLGACID;FVERQAAKYKIGMAAGEYFNAGVLLINL
KKRHDI;FKMA;CEVBEQYKDVMOYQDDILNG;PKGVCYANRFPNFPNDLPMAN
RPAASHDPLVDRITTYTAMPVAVSHYCGAPMWRDCTAWGABERFTBLAGSLTSPER
WRGLAVHVRVFPTRKMLQWRRLISARFLKRY"
1438..2436
/gene="1gta"
/allele="9"
1438..2436
/gene="1gta"
/allele="9"
/note="1gta"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="beta 1,3 N-acetylglucosaminyltransferase"
/protein_id="AAK85138.1"
/db_xref="GI:21654777"
/translation="MQSLVSVLICAVNVEKYFAQSLAAVYNO;TWRNLEILVDDGSTD
GTLIAKDFORDSRIRKILAQNSGLIPSLNIGDLAKSGMGEYIARTDADIAAP
DWIRKIVGEMEKDSITIAMGAMLEVLSEKQGNRLAHHRRGKIWKPTREHDIADP
PFGNDHNNMTIMRSVYDGLRYNTERSDMGOFQYDVYSKGLRVLVYPALVRYTL
HANOVSSVYSRQHEITAGIOKTRANDPFLASMGKTPRDSIEYQIKAVALLEKHL
PEBPERARRFLYCCFRITDLPAGAWLDFAADGRRLFLTRYQFGLIHLRLNKR"
2478..3305
/gene="1gta"
/allele="6"
2478..3305
/gene="1gta"
/allele="6"
/note="1gta"
/codon_start=1
/transl_table=11
/product="beta 1,4 galactosyltransferase"
/protein_id="AAK85139.1"
/db_xref="GI:21654778"
/translation="MGNHVISLASAARRAHADTPRRHGIPFPFPAIMPSELEOA
MAELVPGISAPHYISGVKACFMSHAYLWKQALDEGLYIAVEFDVLLGEGAKFLA
EDAWLQRFDPDSAPIVRLTFMFWHTVLSGVAUYGGRAPFLSHHWGAYTISR
KAMRFLDRPAVLPEGLHPVDLMFDFPDREGVPVCOINPALCAQELHYAKHDON
SALGSLIHDRLNLRKQQRDPSANTPKHRLIRALTKISREKRRQRREGLIIVPQ"
3454..4260
/note="OrfB; allele 1"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK85140.1"
/db_xref="GI:21654779"
/translation="MGNHVISLASAARRAHADTPRRHGIPFPFPAIMPSELEOA
MAELVPGISAPHYISGVKACFMSHAYLWKQALDEGLYIAVEFDVLLGEGAKFLA
EDAWLQRFDPDSAPIVRLTFMFWHTVLSGVAUYGGRAPFLSHHWGAYTISR
KAMRFLDRPAVLPEGLHPVDLMFDFPDREGVPVCOINPALCAQELHYAKHDON
SALGSLIHDRLNLRKQQRDPSANTPKHRLIRALTKISREKRRQRREGLIIVPQ"
4261..4260
/note="OrfB; allele 1"

ORIGIN
Query Match 97.7%; Score 808.8; DB 1; Length 4827;
Best Local Similarity 96.6%; Pred. No. 3.5e-182;
Matches 816; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
Qy 1 ATGCAAAACGCGTATCAGTTAGCTTCGCGCGACAGCGCGACGATTCGCGAT 60
Db 2478 ATGCAAAACGCGTATCAGTTAGCTTCGCGCGACAGCGCGACGATTCGCGAT 2537
Qy 61 ACCTTGGCAGGACGCGATCCGTTCAAGTTTTCAGACGACGATGATGCGGTCTGAAAG 120
Db 2538 ACCTTGGCAGGACGCGATCCGTTCAAGTTTTCAGACGACGATGATGCGGTCTGAAAG 2597

QY	121	CTGGAAACAGGCAATGCGGAACTCGTCCGCCGCTCTGTGGCGCACCCCTATTGGACGGA	180
Db	2598	CTGGAAACAGGCAATGCGGAACTCGTCCGCCGCTCTGTGGCGCACCCCTATTGGACGGA	2657
QY	181	GTGGAAAAAGCCTGCCTTTATGAGCAGCGCCGTAATTGTGGAAAGCAGGCAATTGGACGAAGT	240
Db	2658	GTGGAAAAAGCCTGCCTTTATGAGCAGCGCCGTAATTGTGGAAAGCAGGCAATTGGACGAAGT	2717
QY	241	CTGCCGTATATCAGCGTATTTGAGAGCAGCGTTTACTCGCGGAAGTGAAGAAAAATTC	300
Db	2718	CTGCCGTATATCAGCGTATTTGAGAGCAGCGTTTACTCGCGGAAGGCGGAGAAATTC	2777
QY	301	CTTGTGTGAAGAGCTTTGCTGCAAGAACGCTTTGACCCGGAATACCGCTTTATCGTCCG	360
Db	2778	CTTGTGTGAAGAGCTTTGCTGCAAGAACGCTTTGACCCGGAATACCGCTTTATCGTCCG	2837
QY	361	TTGAAAAAGATGTTTATGACAGTCCGACCTCGGCTCGGCTGGGCGAGTTACGTGCGG	420
Db	2838	TTGAAAAAGATGTTTATGACAGTCCGACCTCGGCTCGGCTGGGCGAGTTACGTGCGG	2897
QY	421	CGCGCTTTCCGCTGTGTGAAAGCAGAACATCGGCGGAGCGCGGAGCTATATCATTTCCGGA	480
Db	2898	CGCGCTTTCCGCTGTGTGAAAGCAGAACATCGGCGGAGCGCGGAGCTATATCATTTCCGGA	2957
QY	481	AAAGCGATGCGGTTTTTCTCTGAGCAGGTTTTGCCTGCGCCGCGGAAGGCGTGACCCC	540
Db	2958	AAAGCGATGCGGTTTTTCTCTGAGCAGGTTTTGCCTGCGCCGCGCGGAAGGCGTGACCCC	3017
QY	541	GTCCATCTGANGATNTTTCAGCGATTTTTTTCAGACAGGGAAGGAATGCCGTTTGCACGTC	600
Db	3018	GTCCATCTGANGATNTTTCAGCGATTTTTTTCAGACAGGGAAGGAATGCCGTTTGCACGTC	3077
QY	601	AATCCCGCTTGTGTGGCCCAAGAGCTGATTAATGCGCAAGTTTCAGACCAAAACAGCGCA	660
Db	3078	AATCCCGCTTGTGTGGCCCAAGAGCTGATTAATGCGCAAGTTTCAGACCAAAACAGCGCA	3137
QY	661	TTGGGACGCTTGATCGAACAGCACCGGCTCTGGAACCGCAAAACAGCAAGGCGGATATCC	720
Db	3138	TTGGGACGCTTGATCGAACAGCACCGGCTCTGGAACCGCAAAACAGCAAGGCGGATATCC	3197
QY	721	CCCGCGAACATTTCAACACCGCTGATCCGCGCTTGACCAAAATAGCAGGGAAGG	780
Db	3198	CCCGCGAACATTTCAACACCGCTGATCCGCGCTTGACCAAAATAGCAGGGAAGG	3257
QY	781	GA AAAAAGCGCGCAAAAGGCGGAACAGTTCAATTGCGCTTCCAAATA	828
Db	3258	GA AAAAAGCGCGCAAAAGGCGGAACAGTTCAATTGCGCTTCCAAATA	3305

RESULT 5	AF470659	3335 bp	DNA	linear	BCT 22-MAY-2002
LOCUS	AF470659				
DEFINITION	Neisseria meningitidis strain M981 1gt-1 locus, complete sequence.				
ACCESSION	AF470659				
VERSION	AF470659.1	GI:21069184			
KEYWORDS					
SOURCE					
ORGANISM	Neisseria meningitidis				
	Neisseria meningitidis				
	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;				
	Neisseriaceae; Neisseria.				
REFERENCE	1 (bases 1 to 3335)				
AUTHORS	Zhu, P., Klutch, M.J., Bash, M.C., Tsang, R.S., Ng, L.K. and Tsai, C.M.				
TITLE	Genetic diversity of three 1gt loci for biosynthesis of				
	lipooligosaccharide (LOS) in Neisseria species				
JOURNAL	Microbiology 148 (Pt 6), 1833-1844 (2002)				
PUBMED	12055303				
REFERENCE	2 (bases 1 to 3335)				
AUTHORS	Zhu, P.				
TITLE	Direct Submission				
JOURNAL	Submitted (22-JUN-2002) DBBAP, CBER/FDA, 8800 Rockville Pike,				
	Bethesda, MD 20892, USA				
FEATURES	Location/Qualifiers				

[illegible]


```

QY 121 CTGGAACAGGCATTCGCGAAGCTGTCCTCCGCGCTTGTGCGCGACCCCTATTGAGCGGA 180
Db 1389 CTGGAACAGGCATTCGCGAAGCTGTCCTCCGCGCTTGTGCGCGACCCCTATTGAGCGGA 1448
QY 181 GTGGAAGAAAGCTGCTTTATGAGCAGCGCGATTTGTGGAAGAGGCAATTGACGAAGGT 240
Db 1449 GTGGAAGAAAGCTGCTTTATGAGCAGCGCGATTTGTGGAAGAGGCAATTGACGAAGGT 1508
QY 241 CTGCGCTATATACCGTATTTGAGAGCAGCGTTTATCTGCGCGAGAGTGAAGAAATTC 300
Db 1509 CTGCGCTATATACCGCTATTTGAGAGCAGCGTTTATCTGCGCGAGAGTGAAGAAATTC 1568
QY 301 CTGCGCGAAGACGCTTGTGCGCAAGACGCTTTGACCCCGGATACCGCTTTATGATCGCG 360
Db 1569 CTGCGCGAAGACGCTTGTGCGCAAGACGCTTTGACCCCGGATTTGATGATCGCG 1628
QY 361 TTGGAACAGATGTTATGCAAGTCTGACCTGCGCTCGCGCGTGGCGGATTAATGCGGG 420
Db 1629 TTGGAACAGATGTTATGCAAGTCTGACCTGCGCTCGCGCGTGGCGGATTAATGCGGG 1688
QY 421 CGGCGCTTCCGCTGTGGAAGAGCACTGGGGGACGGCGGCTATATCATTTCCGA 480
Db 1689 CGGCGCTTCCGCTGTGGAAGAGCACTGGGGGACGGCGGCTATATCATTTCCGA 1748
QY 481 AAAGCATGCGGCTTTTCTGCAAGGTTGCGCGCTCGCGCGAAGGCGTGCACCC 540
Db 1749 AAAGCATGCGGCTTTTCTGCAAGGTTGCGCGCTCGCGCGAAGGCGTGCACCC 1808
QY 541 GTGCAATCATATATGTTGCAAGGATTTTTCAGACGGAAGAAATGCGGTTGGCGAGCTC 600
Db 1809 GTGCAATCATATATGTTGCAAGGATTTTTCAGACGGAAGAAATGCGGTTGGCGAGCTC 1868
QY 601 AATCCGCGCTTGTGCGCGCAAGAGCTGATATATGCCAGTTTTCAGACCAAAACAGCGCA 660
Db 1869 AATCCGCGCTTGTGCGCGCAAGAGCTGATATATGCCAGTTTTCAGACCAAAACAGCGCA 1928
QY 661 TTGGGAGAGGCTATGCAAGCAGCGCGCTCTGGAACCGCAACAGCAAGGCGCATTC 720
Db 1929 TTGGGAGAGGCTATGCAAGCAGCGCGCTCTGGAACCGCAACAGCAAGGCGCATTC 1988
QY 721 CCCGCAACATTCGCAAGCAGCGCGCTGATCGCGCGCTTGAACCAATTCAGAGGGAAG 780
Db 1989 CCCGCAACATTCGCAAGCAGCGCGCTGATCGCGCGCTTGAACCAATTCAGAGGGAAG 2048
QY 781 GAAAAACGCGCGCAAGGCGCGCAAGCTTCATTC 814
Db 2049 GAAAAACGCGCGCAAGGCGCGCAAGCTTCATTC 2082

RESULT 10
AP470663 3246 bp DNA linear BCT 22-MAY-2002
DEFINITION Neisseria meningitidis strain 7880 1gt-1 locus, complete sequence.
ACCESSION AP470663
VERSION AP470663.1 GI:21069199
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 (bases 1 to 3246)
Zhu, P., Klutch, M. J., Bash, M. C., Tsang, R. S., Ng, L. K. and Tsai, C. M.
Genetic diversity of three 1gt loci for biosynthesis of
lipooligosaccharide (LOS) in Neisseria species
Microbiology 148 (Pt 6), 1833-1844 (2002)
JOURNAL
PUBMED
12055303
REFERENCE
2 (bases 1 to 3246)
Zhu, P.
Direct Submission
Submitted (22-JAN-2002) DBPAP, CBER/FDA, 8800 Rockville Pike,
Bethesda, MD 20892, USA
JOURNAL
FEATURES
Location/Qualifiers
1..3246

```

```

/organism="Neisseria meningitidis"
/mol_type="Genomic DNA"
/strain="7880"
/db_xref="taxon:487"
140..1184
/gene="1gtA"
/allele="7"
/pseudo
140..1184
/gene="1gtA"
/allele="7"
/pseudo
140..1184
/gene="1gtA"
/allele="7"
/note="1gtA; glycosyl transferase"
/pseudo
/codon_start=1
/transl_table=11
1226..2065
/gene="1gtB"
/allele="7"
1226..2065
/gene="1gtB"
/allele="7"
/note="1gtB"
/allele="7"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtB"
/protein_id="AAM33872.1"
/db_xref="gi:21069200"
/transl_table="MONHVISLASAARRAHIAATFGARGPFPFDALMPSEALEQA
MAELVPGSAHPDYSIVGEKACFMGSAVLMKQALDEGLYIAVFEDDVLIGSAREPLA
EDAMLOERFSDSAIVLETFMHWLTPSGVADYGRAPLLSEHBMHAGYIISR
KAMRFPLDRFAALPPEGHPVDMWVFSDFPDRGMPVQOLDPALCADELHAKRDON
SALGSLIEHDLRLNRKQRRDPSANTFHLRIALTYISREBERKQREQLIGKLIIV
PFO"
2214..3020
/gene="1gtH"
/allele="8"
2214..3020
/gene="1gtH"
/allele="8"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtH"
/protein_id="AAM33873.1"
/db_xref="gi:21069201"
/transl_table="MONHVISLASAARRAHIAATFGARGPFPFDALMPSEALEQA
MAELVPGSAHPDYSIVGEKACFMGSAVLMKQALDEGLYIAVFEDDVLIGSAREPLA
EDAMLOERFSDSAIVLETFMHWLTPSGVADYGRAPLLSEHBMHAGYIISR
KAMRFPLDRFAALPPEGHPVDMWVFSDFPDRGMPVQOLDPALCADELHAKRDON
SALGSLIEHDLRLNRKQRRDPSANTFHLRIALTYISREBERKQREQLIGKLIIV
PFO"
93.7%; Score 775.6; DB 1; Length 3246;
Best Local Similarity 97.1%; Pred. No. 2.9e-174;
Matches 790; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```

Db 1466 CTCGCCGATATCCGCCGATTTGAGACGAGTTTACTCCGCCGAAAGCGCGAGAAATTC 1525
Qy 301 CTGGCGAAGACGCTGGCTGCAAGAACGCTTTGACCCCGATACCGCTTTATCGTCCG 360
Db 1526 CTGGCGAAGACGCTGGCTGCAAGAACGCTTTGACCCCGATACCGCTTTATCGTCCG 1585
Qy 361 TTGAAAACGATGTTTATGACAGTCTGACCTCGCCCTCCGCGCTGCGAGATTACTGCGG 420
Db 1586 TTGAAAACGATGTTTATGACAGTCTGACCTCGCCCTCCGCGCTGCGAGATTACTGCGG 1645
Qy 421 CGGCGCTTCCGCTGTTGAAAGCGAACACTGGGGGACGCGCGGCTATATCATTTCCCGA 480
Db 1646 CGGCGCTTCCGCTGTTGAAAGCGAACACTGGGGGACGCGCGGCTATATCATTTCCCGA 1705
Qy 481 AAGGCGATGCGGTTTCTTCTGACAGGTTTCCGCGCTCCGCGCGCGGAGGCTGACCCC 540
Db 1706 AAGGCGATGCGGTTTCTTCTGACAGGTTTCCGCGCTCCGCGCGCGGAGGCTGACCCC 1765
Qy 541 GTGATCTGATGATGTTTCAAGCGATTTTTCGACAGGGAAGGAATGCGGTTTCCGACGTC 600
Db 1766 GTGATCTGATGATGTTTCAAGCGATTTTTCGACAGGGAAGGAATGCGGTTTCCGACGTC 1825
Qy 601 AATCCGCGCTTGTGCGCCCAAGAGCTGCAATTATGCCAAGTTTCAAGCAAGCAAGCGCA 660
Db 1826 AATCCGCGCTTGTGCGCCCAAGAGCTGCAATTATGCCAAGTTTCAAGCAAGCAAGCGCA 1885
Qy 661 TTGGGAGCGCTGATTCGAACAGCAGCGCTCTCTGTAACCGCAAGCAAGCAAGCGCAATTC 720
Db 1886 TTGGGAGCGCTGATTCGAACAGCAGCGCTCTCTGTAACCGCAAGCAAGCAAGCGCAATTC 1945
Qy 721 CCCGCAACACATTCGAACAGCAGCGCTGATCCGCGCTTGAACCAATTCAGAGGGAAGG 780
Db 1946 CCCGCAACACATTCGAACAGCAGCGCTGATCCGCGCTTGAACCAATTCAGAGGGAAGG 2005
Qy 781 GAAAAAGCGCGCAAGGCGCGCAACAGTTCAATTG 814
Db 2006 GAAAAAGCGCGCAAGGCGCGCAACAGTTCAATTG 2039

RESULT 11
AP470662
LOCUS AP470662 3289 bp DNA linear BCT 22-MAY-2002
DEFINITION Neisseria meningitidis strain M120 1gt-1 locus, complete sequence.
ACCESSION AP470662
VERSION AP470662.1 GI:21069195
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
1 (bases 1 to 3289)
Zhu, P., Klutcher, M.J., Baeh, M.C., Tsang, R.S., Ng, L.K. and Tsai, C.M.
Genetic diversity of three 1gt loci for biosynthesis of
Microisogenescharide (LOS) in Neisseria species
Microbiology 148 (Pt 6), 1833-1844 (2002)
12055303
2 (bases 1 to 3289)
Zhu, P.
Direct Submission
Submitted (22-JAN-2002) DBPAP, CBBR/FDA, 8800 Rockville Pike,
Bethesda, MD 20892, USA
FEATURES
source
1..3289
/location="Qualifiers"
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="M120"
/db_xref="taxon:487"
106..1152
/gene="1gtA"
106..1152
/allele="6"
f./gene="1gtA"

CDS
106..1152
f./gene="1gtA"

/allele="6"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtA"
/protein_id="AA033869.1"
/db_xref="GI:21069196"
/translation="MPSEAFRHRRAVREKLOPLVSVLCARVNEYKFPAGSLAAVNO
TMCNDLIVDGSIDGTALIAIDPQKRSRIKILAOQNSGLIPSLNGLBELASG
GSEYIARTADIAIPDWITEKIVGEMKDRSITAMGAMLEVLSEBQKRLAHHHG
KIWKETREDIADFPFGNPIHNNMTIMRRESVIDGRLNTEREDWAEYQFVYSK
LRLAYPEALVKEVRLHANOVSKYSVRQHEIAOGIOKTRADPLOSMEFKRFDLB
VROIRAVAYELIEKLIKPEBDFBARAFLYQCKIKRTDITLPAGMLDPAADGRRLFTL
ROYFGIIRHLNKR"
1194..2033
/gene="1gtB"
/allele="7"
1194..2033
/gene="1gtB"
/allele="7"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtB"
/protein_id="AA033870.1"
/db_xref="GI:21069197"
/translation="MKNHVISLASAABERRAHIAATFGARCIPEQFPDAMPSEBLEQA
MAELVGLSAPHYLSGVEKACFMSHVAIVMKQALDGLPIAVFEDVILGSEAKFLA
EDAMLOERFDKDSAFIVRLTFMFMHVLSPSGVADYGRAPFLSEHMGTAGYIISR
KAMRFFLRPAVLPEERLKAVDMMFLSSFLDKGGMVTCOLPDLCAOELHVAKPHON
SALGSLIHDRILNKKQRRDSPATFPHRLIRALTKISREBKRRQREQLGKIIV
PQ"
2182..2988
/gene="1gtH"
/allele="7"
2182..2988
/gene="1gtH"
/allele="7"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="1gtH"
/protein_id="AA033871.1"
/db_xref="GI:21069198"
/translation="MKNHVISLASAABERRAHIAATFGARCIPEQFPDAMPSEBLEQA
MAELVGLSAPHYLSGVEKACFMSHVAIVMKQALDGLPIAVFEDVILGSEAKFLA
EDAMLOERFDKDSAFIVRLTFMFMHVLSPSGVADYGRAPFLSEHMGTAGYIISR
KAMRFFLRPAVLPEERLKAVDMMFLSSFLDKGGMVTCOLPDLCAOELHVAKPHON
SALGSLIHDRILNKKQRRDSPATFPHRLIRALTKISREBKRRQREQLGKIIV
PQ"
93.7%; Score 775.6; DB 1; Length 3289;
Best Local Similarity 97.1%; Pred. No. 2.9e-174;
Matches 790; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 1 ATGCAAAACACGTTATCACTTACCTCCGCGCAAGACGAGGCGCGCACTTCCCAT 60
Db 1194 ATGCAAAACACGTTATCACTTACCTCCGCGCAAGACGAGGCGCGCACTTCCCAT 1253
Qy 61 ACCTTCGAGGCAAGGATCCCGTTTTCAGAGCACTGATGCGGCTGAAAGG 120
Db 1254 ACCTTCGAGGCGGCGGATCCCGTTTTCAGAGCACTGATGCGGCTGAAAGG 1313
Qy 121 CTGGAACAGGCAATGCGGAATCGTCCCGCTGTGCGGCAACCCCTATTGAGCGGA 180
Db 1314 CTGGAACAGGCAATGCGGAATCGTCCCGCTGTGCGGCAACCCCTATTGAGCGGA 1373
Qy 181 GTGAAAAAGCGCTGTTTATGAGCCAGCGCGTATTTGGAAGCAGGCAATTGACGAGCT 240
Db 1374 GTGAAAAAGCGCTGTTTATGAGCCAGCGCGTATTTGGAAGCAGGCAATTGACGAGCT 1433
Qy 241 CTGCGGTATACCGGTATTTGAGAGCAGCTTTTACTCGGCGAAGGTGAGAAATTC 300

Db 1434 CTGCCGTATATGCGCGTATTTGAGAGCAGCTTTTCTCGCGAAGCGCGGAGAAATTC 1493
Qy 301 CTTCGCGAAGACGCTGGCTGCAAGAACGCTTTGACCGCGATACCGCTTTATCGTCCG 360
Db 1494 CTTCGCGAAGACGCTGGCTGCAAGAGCGTTTGAACAAGATTCGCCCTTTATCGTCCG 1553
Qy 361 TTGGAACGATGTTTATGACGCTGACCTGCGCTCGCGGCTGCGGATTAATCGCGGG 420
Db 1554 TTGGAACGATGTTTATGACGCTGACCTGCGCTCGCGGCTGCGGATTAATCGCGGG 1613
Qy 421 CGGCGCTTTCCGCTGTTGGAAGACGAACTGCGGGAACGCGCGGCTTATCAATTTCCGA 480
Db 1614 CGGCGCTTTCCGCTGTTGGAAGACGAACTGCGGGAACGCGCGGCTTATCAATTTCCGA 1673
Qy 481 AAGAGCATGCGGTTTTTCTCGGACAGGTTTGGCGCGCTCGCGCGAAGGCGTGCACCGC 540
Db 1674 AAGAGCATGCGGTTTTTCTCGGACAGGTTTGGCGCGCTCGCGCGAAGGCGTGCACCGC 1733
Qy 541 GTGCAATGATGATGTTTCAAGCGATTTTTCGACAGGGAAGAAATGCGGTTTGCAGCTC 600
Db 1734 GTGCAATGATGATGTTTCAAGCGATTTTTCGACAGGGAAGAAATGCGGTTTGCAGCTC 1793
Qy 601 AATCCGCGCTGTCGCGCGCAAGAGCTGCATTAATGCAAGTTTCAAGCAAAACAGCGCA 660
Db 1794 GATCCGCGCTTGTGCGCGCAAGAGCTGCATTAATGCAAGTTTCAAGCAAAACAGCGCA 1853
Qy 661 TTGGGAGCGCTGATTCGACAGCAGCGCTCTCTGAAACCGGAAACAGCAAGGCGCATTC 720
Db 1854 TTGGGAGCGCTGATTCGACAGCAGCGCTCTCTGAAACCGGAAACAGCAAGGCGCATTC 1913
Qy 721 CCGCGCAACATTCGAAACAGCGCGCTGATTCGCGCGCTTGAACCAATTCAGCAGGAAAG 780
Db 1914 CCGCGCAACATTCGAAACAGCGCGCTGATTCGCGCGCTTGAACCAATTCAGCAGGAAAG 1973
Qy 781 GAAAAAGCGCGCAAGGCGCGCAAGCTTCATTG 814
Db 1974 GAAAAAGCGCGCAAGGCGCGCAAGCTTCATTG 2007

RESULT 12
AF470655
LOCUS AF470655 3383 bp DNA linear BCT 22-MAY-2002
DEFINITION Neisseria meningitidis strain 35E 1gt-1 locus, complete sequence.
ACCESSION AF470655
VERSION AF470655.1 GI:21069168
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
1 (bases 1 to 3383)
Zhu, P., Klutch, M. J., Baah, M. C., Tsang, R. S., Ng, L. K. and Tsai, C. M.
Genetic diversity of three 1gt loci for biosynthesis of
lipooligosaccharide (LOS) in Neisseria species
Microbiology 148 (Pt 6), 1833-1844 (2002)
JOURNAL
PUBMED 12055304
REFERENCES
AUTHORS
TITLE
JOURNAL
Submitted (22-JAN-2002) DBPAP, CBBR/FDA, 8800 Rockville Pike,
Bethesda, MD 20892, USA
FEATURES
source
1..3383
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="35E"
/db_xref="taxon:487"
140..1186
/gene="1gtA"
/allele="16"
140..1186
/gene="1gtA"
/allele="16"
140..1186
/gene="1gtA"
/allele="16"

/note="glycosyl transferase"
/codon_start=1
/transl_table=1
/product="1gtA"
/protein_id="AA033849.1"
/db_xref="gi:21069168"
/translation="MPSRAFRHRAHRENKIQLIVSVLICAYNEKYPAQSLAAVNO
TWCNDILITVDGSDGTALIAQRFQBOQDRIKITLAQKQNSGLIPSLTGLDELAKSG
MGEYIARTDADIDIAAPDWITEKI VGEEMEDRSI IAMGAWLEVLSEKQGNRLARHHEG
KIWKPPTHEDIAQFPFGNDIHNNTMERSVTDGGLRYNTERDMADYQFWYDYSK
LGRLLAYPEALVETKRLHAMOVSSKYSIKOHEIAQGI OKTANDEFLQSWGFTKRDLSL
YROIKAVALYELKEHLPEDEPERARRLFYQCFKRTDPLPAGVMDPANGMRRLFTL
ROYGILRLIKON"
1228..2067
/gene="1gtB"
/allele="4"
1228..2067
/gene="1gtB"
/allele="4"
/note="glycosyl transferase"
/codon_start=1
/transl_table=1
/product="1gtB"
/protein_id="AA033850.1"
/db_xref="gi:21069170"
/translation="MONHVISLAAABRAHIAATPGRHGIPQFPDMLPSEBLEOA
MAELVPGLSAPYLSGVEKACFMGHVVLIMKQALDEGLPYLAIVEDDVLGGAQPLA
EDTLBERFPDPSAFIVRLBTMPAKVIVRPKVLNTERSPFLSEHWGATGISR
EAMKFLERFALVPAEWIKAVDMMFTFYFDKEMPYQVNPALCTOELHAYAKFLSKN
SMLSDEKDEORRRHRSIKVMFDIKRALGKGRBKRMERORQAELEKAGRV
ISFY"
2216..3058
/gene="1gtB"
/allele="4"
2216..3058
/gene="1gtB"
/allele="4"
/note="glycosyl transferase"
/codon_start=1
/transl_table=1
/product="1gtB"
/protein_id="AA033851.1"
/db_xref="gi:21069171"
/translation="MONHVISLAAABRAHIAATPGRHGIPQFPDMLPSEBLEOA
MAELVPGLSAPYLSGVEKACFMGHVVLIMKQALDEGLPYLAIVEDDVLGGAQPLA
EDTLBERFPDPSAFIVRLBTMPAKVIVRPKVLNTERSPFLSEHWGATGISR
EAMKFLERFALVPAEWIKAVDMMFTFYFDKEMPYQVNPALCTOELHAYAKFLSKN
SMLSDEKDEORRRHRSIKVMFDIKRALGKGRBKRMERORQAELEKAGRV
ISFY"
ORIGIN
Query Match 91.0%; Score 753.2; DB 1; Length 3383;
Best Local Similarity 95.3%; Pred. No. 6.4e-169;
Matches 776; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
Qy 1 ATGCAAAACCAAGTATAGCTTAGCTTCGCGCAGAACGCAAGGCGCAATTCGCGAT 60
Db 1228 ATGCAAAACCAAGTATAGCTTAGCTTCGCGCAGAACGCAAGGCGCAATTCGCGCA 1287
Qy 61 ACCTTCGCGAGCAGCGATCCGTTTCAAGTTTTCGACGCACTGATGCCGTGGAAG 120
Db 1288 ACCTTCGCGAGCAGCGATCCGTTTCAAGTTTTCGACGCACTGATGCCGTGGAAG 1347
Qy 121 CTGGAACAGGCAATGCGGCACTGCTCCCGCTTGTTCGCGCACCCCTATTGAGCGGA 180
Db 1348 CTGGAACAGGCAATGCGGCACTGCTCCCGCTTGTTCGCGCACCCCTATTGAGCGGA 1407
Qy 181 GTGGAAGAAAGCTCTTTATGAGCGCAGCGCTATTGGAAGCAGCAATTCGGAAGGT 240
Db 1408 GTGGAAGAAAGCTCTTTATGAGCGCAGCGCTATTGGAAGCAGCAATTCGGAAGGT 1467
Qy 241 CTGCCGTATATCAGCGTATTTGAGAGCAGCTTTTACCTCGCGAAGGTGAGGAATAATTC 300

QY 301 CTGCGAAGACGCTTGGCTGCAAGAACGCTTTGACCCGAGTACCGCTTTATGCTCCG 360
Db 1494 CTGGCGGAAGACGCTTGGCTGCAAGAACGCTTTGACCCGAGTACCGCTTTATGCTCCG 1553
QY 361 TTGGAAACGATGTTTATGCAAGTCTGACCTCGCTCCGCGTGGCGGATTACTCGCGG 420
Db 1554 TTGGAAACGATGTTTATGCAAGTCTGACCTCGCTCCGCGTGGCGGATTACTCGCGG 1613
QY 421 CGGCGCTTCCGCTGTTGGAAAGCAACCTGGGGGACCGCGGCTATATCATTTCCGCA 480
Db 1614 CGGCGCTTCCGCTGTTGGAAAGCAACCTGGGGGACCGCGGCTATATCATTTCCGCA 1673
QY 481 AAAAGCATGCGATTTTCTCTGACAGATTTGCGCCCTGCGCCGCAAGGCTGCACCCC 540
Db 1674 AAGGCGATGCGATTTTCTCTGACAGATTTGCGCCCTGCGCCGCTGCAAGGCTGCACCCC 1733
QY 541 GTGCACTGATGATGTTTCAGCCGATTTTTCGACAGGAAAGATGCGGTTTCCGACGCTC 600
Db 1734 GTGCACTGATGATGTTTCAGCCGATTTTTCGACAGGAAAGATGCGGTTTCCGACGCTC 1793
QY 601 AATCCCGCTTGTGCGCCGCAAGAGTGCATTTATGCCAAGTTTCAAGCAAAACAGCGCA 660
Db 1794 AATCCCGCTTGTGCGCCGCAAGAGTGCATTTATGCCAAGTTTCAAGCAAAACAGCGCA 1853
QY 661 TTGGGAGGCTGATGCAACACGCGCTCTCTGAACCGCAAAAGCGCGATTC 720
Db 1854 TTGGGAGGCTGATGCAACACGCGCTCTCTGAACCGCAAAAGCGCGATTC 1913
QY 721 CCGCGCAACATTCGAAACACGCGCTGATCCGCGCTTGAACCAAAATCGAGCGGAAAG 780
Db 1914 CCGCGCAACATTCGAAACACGCGCTGATCCGCGCTTGAACCAAAATCGAGCGGAAAG 1973
QY 781 GAAAAACGCGCGCAAGGCGCGCAACGATTCATG 814
Db 1974 GAAAAACGCGCGCAAGGCGCGCAACGATTCATG 2007

RESULT 14
LOCUS AY134878 3066 bp DNA linear BCT 02-AUG-2003
DEFINITION Neisseria polysaccharea strain 87043 lgt-1 locus, complete
ACCESSION AY134878
VERSION AY134878.1 GI:33413664
KEYWORDS
SOURCE Neisseria polysaccharea
ORGANISM Neisseria polysaccharea
Neisseria polysaccharea
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 3066)
AUTHORS Zhu, P. and Tsai, C.M.
TITLE Galactosyltransferase gene family in Neisseria: Implication for
understanding the evolution of prokaryotic enzymes responsible for
glycosylation
JOURNAL Unpublished
REFERENCES 2 (bases 1 to 3066)
AUTHORS Zhu, P.
TITLE Direct Submission
JOURNAL Submitted (22-JUN-2002) LBP/DBPAP/CBER, FDA, 8800 Rockville Pike,
Bethesda, MD 20892, USA
FEATURES
source location/Qualifiers
1..3066
/organism="Neisseria polysaccharea"
/mol_type="genomic DNA"
/db_xref="taxon:489"
/note="strain 87043"
181..1230
/gene="lgtA"
181..1230
/allele="19"
181..1230
/gene="lgtA"
/allele="19"

CDS
181..1230
/gene="lgtA"
/allele="19"

/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="lgtA"
/protein_id="AA08517.1"
/db_xref="GI:33413665"
/translation="MPSEAFRRHRAVRENKLOPIVSVLICAVNEKYPAGSLAAVNO
TWRNDILIVDDGTDGLAIAPRQEDDGGIRILAOBRNSGLIPSLNIGDLAKSG
GGGIYARTDADIDIAAPDWIKETIVGEMKDSIIAMGAWLYLSERKGNLARHEH
GKIRKAPTRHEDIAAPFPFNPINHTNITMRSTYDGLRINTEDMADYQPTVDV
KUGRLAYPEALVKTRLHANOVSSKYSVRQHEIAOGIQTANRPFLOSGMGRTERPDSL
EYRQIKAVAYBELHKLPEBDPERARRELQCFKRTDTPAGAWLDFADGRMRLFT
LROYGILRLRLKRR"
1272..2111
/gene="lgtB"
/allele="19"
1272..2111
/gene="lgtB"
/allele="19"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="lgtB"
/protein_id="AA08518.1"
/db_xref="GI:33413666"
/translation="MONHVISLAAERRAHIAITFGRHGIPOFPDAMPSELEQA
MAELVPGISAPHYLSGVKACFMSTHTVLMKQALDESLPIYIAVFEDVILGGAETFLA
EDGMVKERPDSDAFIVRLFTFMHVLVPSGVADYCGRAPPLLSBHPGAGYIISR
KAMRPFIDRFALPSEBGLPYDMMWFGPDDBGMPYCOLNAPLCAOELIYAKPHDN
SALGSLIHDLRLNRKQORRDS PANTFHRILRLTKI SRERKRRQRBQLIGKTIY
PQ"
2260..3066
/gene="lgtH"
/allele="12"
2260..3066
/gene="lgtH"
/allele="12"
/note="glycosyl transferase"
/codon_start=1
/transl_table=11
/product="lgtH"
/protein_id="AA08519.1"
/db_xref="GI:33413667"
/translation="MONHVISLAAERRAHIAITFGRHGIPOFPDAMPSELEQA
MAELVPGISAPHYLSGVKACFMSTHTVLMKQALDESLPIYIAVFEDVILGGAETFLA
EDTMLEKFPDSDAFIVRLFTFMHVLVPSGVADYCGRAPPLLSBHPGAGYIISR
KAMRPFIDRFALPSEBGLPYDMMWFGPDDBGMPYCOLNAPLCAOELIYAKPHDN
ROESYRNRSFKVLKRALGKIGREIERARERKQKLEKILGRHVDFE"

ORIGIN
Query Match 90.4%; Score 748.4; DB 1; Length 3066;
Beet Local Similarity 95.0%; Pred. No. 9e-168;
Matches 773; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 1 ATGCAAAACGAGTTATGAGCTTCCGCGCAAGCAGGCGCACATTTGCCAT 60
Db 1272 ATGCAAAACGAGTTATGAGCTTCCGCGCAAGCAGGCGCACATTTGCCAT 1331
QY 61 ACCTTGGCAGGACGCGATCCCGTTTCACTGATGATGCGCTGAAAGG 120
Db 1332 ACCTTGGCAGGACGCGATCCCGTTTCACTGATGATGCGCTGAAAGG 1391
QY 121 CTGGAACAGGCAATGGCGGAAGTCCGCGCTTGTGGCGCACCCCTATTGAGCGGA 180
Db 1392 CTGGAACAGGCAATGGCGGAAGTCCGCGCTTGTGGCGCACCCCTATTGAGCGGA 1451
QY 181 GTGAAAAAGCTCTGTTATGAGCGACGCGTATTTGGAAGCAGGATTGACGAAGGT 240
Db 1452 GTGAAAAAGCTCTGTTATGAGCGACACCGTATTTGGAAGCAGGCGCTGACGAAGGT 1511
QY 241 CTGCGGTATCACCGTATTTGAGGACGATTTTACCTGGCGAAGGTGAGGAAAAATTC 300
Db 1512 TTACCGTATTTCCGCTATTTGAGGACGATTTTACCTGGCGAAGGTGAGGAAAAATTC 1571

FEATURES	Source
Db	301 CTGCGCAAGACGCTTGCGTCGCAAGAACGCTTTGACCCGGATACCGCTTTATGCTCCG 360
Qy	1572 CTTCGCGAAGACGCGCTGCTGTAAGAACGCGTTGACCCGGATTCGCGCTTTATCGTTCGT 1631
Db	361 TTGGAACGATGTTTATGACGCTGCTGACCTTCGCGCGTTCGCGGATTAATCTCGCG 420
Qy	1632 TTGGAACGATGTTTATGACGCTGCTGACCTTCGCGCGTTCGCGGATTAATCTCGCG 1691
Db	421 CGCGGCTTTCCGCTGTTGGAAAGCGAACACATGCGGGGACGCGCGGCTATATCATTTTCCCGA 480
Qy	1692 CGTGGCTTTCCGCTGTTGGAAAGCGAACACATGCGGGGACGCGCGGCTATATCATTTTCCCGA 1751
Db	481 AAAGCGATCGGCTTTTTCCTGGACAGGTTTTCGCGCTTCGCGCGCGCGGATGACACCC 540
Qy	1752 AAAGCGATCGGCTTTTTCCTGGACAGGTTTTCGCGCTTCGCGCGCGCGGATGACACCC 1811
Db	541 GTCGATCTGATGATGTTTTCGACGATTTTTCGACAGGAAAGAAATGCGGTTTTCGACGCTC 600
Qy	1812 GTCGATCTGATGATGTTTTCGACGATTTTTCGACAGGAAAGAAATGCGGTTTTCGACGCTC 1871
Db	601 AATCCCGCTTTGCGGCGCCCAAGGCTGCTATATGCGCAAGTTTTCAGACACCAAAACGCGCA 660
Qy	1872 AATCCCGCTTTGCGGCGCCCAAGGCTGCTATATGCGCAAGTTTTCAGACACCAAAACGCGCA 1931
Db	661 TTGGGACGCTGATGACGAACGACGCGCTTCGGAACGCGGAAACAGCAAAAGGCGCATTC 720
Qy	1932 TTGGGACGCTGATGACGAACGACGCGCTTCGGAACGCGGAAACAGCAAAAGGCGCATTC 1991
Db	721 CCCGCAACACATTCAAACACCGCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG 780
Qy	1992 CCCGCAACACATTCAAACACCGCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG 2051
Db	781 GAAAAACGCGCGCAAAAGCGCGCAACGTTAATCG 814
Qy	2052 GAAAAACGCGCGCAAAAGCGCGCAACGTTAATCG 2085
RESULT 15	
LOCUS	NMA22491
DEFINITION	Neisseria meningitidis serogroup A strain Z2491 complete genome; segment 2/7.
ACCESSION	AL162753
VERSION	AL157959
KEYWORDS	AL162753.2 GI:7379120
SOURCE	Neisseria meningitidis Z2491
ORGANISM	Neisseria meningitidis Z2491
REFERENCE	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.
AUTHORS	1 (bases 1 to 349061) Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Koresli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,K., Feltham,T., Hamlin,N., Holtroft,S., Jagers,K., Leach,K., Moulton,S., Mungall,K., Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M., Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
TITLE	Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
JOURNAL	Nature 404 (6777), 502-506 (2000)
PUBMED	10761919
REFERENCE	2 (bases 1 to 349061)
AUTHORS	Parkhill,J.
TITLE	Direct Submission
JOURNAL	Submitted (30-MAR-2000) Submitted on behalf of the Neisseria sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT	Notes: Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web. (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES	Location/Qualifiers
Source	f...349061

```

repeat_unit
    /organism="Neisseria meningitidis Z2491"
    /mol_type="genomic DNA"
    /strain="Z2491"
    /db_xref="taxon:122587"
    /note="serogroup: A"
    complement(24..206)
    /note="NTR repeat; hmms hit to HMM ATR (1 - 183), score:
    310.39"
    label=ATR
RBS
    209..212
gene
    220..681
    /gene="NMA0368"
CDS
    220..681
    /gene="NMA0368"
    /note="NMA0368, probable integral membrane protein, len:
    153 aa; contains four probable transmembrane domains"
    codon_start=1
    /transl_table=11
    product="putative integral membrane protein"
    protein_id="CAB83669.1"
    db_xref="GI:7379121"
    /db_xref="UniProt/TxEmbl:Q9JWH7"
    translation="MOBONRKSSPPIVMLAVLVAISLVNVAFLYLNHNSNEGTLV
    LIIISPLSLDIRCAVAYANVTALIVLAIIRKKVVPAARFGLATVAFSKAAVTV
    DEAGTSDIIVRYGGAFILWTAAPFAVASIGIFAGNNKRRKAASADGINKDV"
    671..1492
    /gene="hemk"
    671..1492
    /gene="hemk"
CDS
    671..1492
    /gene="hemk"
    /note="NMA0369, hemk, Hemk protein, len: 273 aa; similar
    to e.g. HEHK_ECOLI P37186 HEHK protein (277 aa), fasta
    scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
    PS00092 N-6 Adenine-specific DNA methylases signature"
    codon_start=1
    /transl_table=11
    product="Hemk protein"
    protein_id="CAB83670.1"
    db_xref="GI:7379122"
    /db_xref="GOA:Q9JWH6"
    /db_xref="InterPro:IPR002052"
    /db_xref="InterPro:IPR002052"
    /db_xref="InterPro:IPR004556"
    /db_xref="UniProt/TxEmbl:Q9JWH6"
    translation="MTFDKM/LGSKLPRKNARMLLYOVSEYRVQLITRGGEPMDEV
    RQRDRLAKORLRINPEVAYLIGAREFYGRFTVPVSVLIPRTTEHLVEAVLARLPEN
    GRVMVLIQSGSAVVALEERPDFAFVASISIPPALETARNADICARVEFAYGSWF
    DTMSDESKMIIIVSNPYILENGDKHSOGLDRREPOLITDPSDGISCRIITAQAP
    DRLAEGRILLRHGFDDGAARVGVLAEANGFSGVETLPDLAIGLDREVLTGKTKLK"
    831..840
misc_feature
    831..840
    /gene="hemk"
    /note="Core DNA uptake sequence: gccgctcgaa"
    label=DUS
misc_feature
    1162..1171
    /gene="hemk"
    /note="Core DNA uptake sequence: gccgctcgaa"
    label=DUS
misc_feature
    1187..1207
    /gene="hemk"
    /note="PS00092 N-6 Adenine-specific DNA methylases
    signature"
    complement(1279..1288)
misc_feature
    1288
    /note="Core DNA uptake sequence: gccgctcgaa"
    label=DUS
RBS
    1565..1568
gene
    1578..2966
    /gene="NMA0370"
    1578..2966
    /gene="NMA0370"
    /note="NMA0370, probable integral membrane protein, len:
    462 aa; similar to hypothetical proteins e.g. Y325_HABIN
    P4460 hypothetical protein HI0325 (450 aa), fasta_scores;
    E(): 0, 49.8% identical in 464 aa overlap"
    codon_start=1

```


QY 721 CCCGCCAACAATTCAAAACCCGCTGATCGCGCTTGA CCAAAATCAGCAGGAAAG 780
Db 163504 CCCGCCAACAATTCAAAACCCGCTGATCGCGCTTGA CCAAAATCAGCAGGAAAG 163563
QY 781 GAAAAAGCCGCCAAGGCGCAACAGTTCAATTG 814
Db 163564 GAAAAAGCTCGGCAAGGCGCAACAGTTAATCG 163597

Search completed: April 7, 2006, 18:04:58
Job time : 5663.11 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 15:58:06 ; Search time 933.991 Seconds
(without alignments)
5908.370 Million cell updates/sec

Title: US-09-211-691-1

Perfect score: 828
Sequence: 1 atgcaaacacagctatcacg.....tcattgtgccttccataaa 828

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	828	100.0	828	2	AAx84281 N. mening
2	826.4	99.8	46593	3	AAx81456 N. mening
3	826.4	99.8	110000	3	AAx81459_5
4	826.4	99.8	349980	3	AAx81461
5	726	87.7	5859	2	AAx81461
6	726	87.7	5859	2	AAx81461
7	721.2	87.1	837	10	ABx38360 N. gonorr
8	471.8	57.0	840	10	ABx38368 N. gonorr
9	426	51.4	840	8	ABx09919 N. mening
10	93	11.2	14547	13	ADx05534 Haemophil
11	93	11.2	10645	13	ADx05534 Haemophil
12	86.6	10.5	110000	2	AAx82063_05
13	73.2	8.8	768	12	ADx02801 DNA encod
14	73.2	8.8	4583	14	ADx09334 M. catarr
15	73.2	8.8	5584	14	ADx09335 M. catarr
16	73.2	8.8	65792	4	AAx28544 Genomic f
17	71.4	8.6	5570	14	ADx09333 M. catarr
18	57	6.9	3625	13	ADx05481 Haemophil
19	57	6.9	110000	2	AAx82063_08

20	57	6.9	135356	13	ADx05646	Adx05646 Haemophil
21	48.2	5.8	2000	8	ADx71938	Adx71938 Rice gene
22	47	5.7	2000	8	ADx71938	Adx71938 Rice gene
23	39.8	4.8	4532	4	ABx10847	Abx10847 Drosophila
24	39.4	4.8	4455	14	ACx70656	Acx70656 M. xanthu
25	39.4	4.8	4455	14	ACx70656	Acx70656 M. xanthu
26	39.4	4.8	4455	14	ACx70656	Acx70656 M. xanthu
27	39	4.7	399	14	ACx66474	Acx66474 M. xanthu
28	39	4.7	399	14	ACx66474	Acx66474 M. xanthu
29	39	4.7	399	14	ACx66474	Acx66474 M. xanthu
30	38.6	4.7	753	11	ABx11686	Abx11686 Pseudomon
31	38.6	4.7	954	13	ADx42128	Adx42128 Bacteri
32	38.6	4.7	1260	11	ABx11889	Abx11889 Pseudomon
33	38.6	4.7	1635	11	ABx11889	Abx11889 Pseudomon
34	38.6	4.7	1743	11	ABx12070	Abx12070 Pseudomon
35	38	4.6	1760	14	ABx26876	Abx26876 P. rad
36	37.8	4.6	2611	6	ABx56602	Abx56602 S. macro
37	37.8	4.6	4540	6	ABx56604	Abx56604 S. macro
38	37.4	4.5	1329	13	ADx20087	Adx20087 Plant CN
39	36.8	4.4	31422	3	AAx92302	Aax92302 S. averm
40	36.8	4.4	31422	4	AAx92302	Aax92302 S. averm
41	36.8	4.4	39912	14	ADx43464	Adx43464 S. averm
42	36.4	4.4	1122	8	ACA44178	Aca44178 Prokaryot
43	36.4	4.4	7788	10	ADx23894	Adx23894 Streptomy
44	36.4	4.4	19211	3	AAx81507	Aax81507 N. mening
45	36.4	4.4	37360	10	ADx23892	Adx23892 Streptomy

ALIGNMENTS

RESULT 1	
AAx84281	standard; DNA; 828 BP.
ID	AAx84281
XX	
AC	AAx84281;
XX	
DT	08-SEP-1999 (first entry)
XX	
DE	N. meningitidis Beta-1,4-galactosyltransferase coding sequence.
XX	
KW	Beta-1,4-galactosyltransferase; 1978; fusion protein; catalytic domain;
KW	glycosyltransferase; accessory enzyme; nucleotide sugar formation;
KW	saccharide donor; oligosaccharide synthesis;
KW	carbohydrate structure development; ds.
XX	
OS	Neisseria meningitidis.
XX	
PN	W09931224-A2.
XX	
PD	24-JUN-1999.
XX	
PF	15-DEC-1998; 98WC-CA001180.
XX	
PR	15-DEC-1997; 97US-0069443P.
XX	
PR	14-DEC-1998; 98US-00211691.
XX	
PA	(CANADA) NAT RES COUNCIL CANADA.
XX	
PI	Gilbert M, Young NM, Wakarchuk WM;
XX	
DR	WPI, 1999-395174/33.
XX	
PT	F-PSDB; AAY22156.
XX	
PS	A new glycosyltransferase fusion protein useful in the enzymatic
XX	synthesis of oligosaccharides.
XX	
PS	Example 2; Fig 2; 63pp; English.
XX	
CC	This sequence encodes the Neisseria meningitidis Beta-1,4-
CC	galactosyltransferase (also referred to as 1978). The invention relates
CC	to a nucleic acid encoding a fusion protein that comprises a
CC	glycosyltransferase catalytic domain and a catalytic domain from an

accessory enzyme that is involved in formation of a nucleotide sugar
 CC which is a saccharide donor for a glycosyltransferase reaction. The
 CC fusion protein is useful in the enzymatic synthesis of oligosaccharides.
 CC The fusion proteins are able to catalyze more than one reaction involved
 CC in the enzymatic synthesis. This is useful for the development of
 CC therapeutic agents that have specific carbohydrate structures.
 CC Carbohydrates are involved in recognition elements on the surface of
 CC cells. The fusion protein can be used for the synthesis of both natural
 CC carbohydrates and synthetic derivatives with novel properties. The fusion
 CC polypeptide allows two glycosyltransferase reactions in a single vessel,
 CC provides improved yields of end products. Additionally, cleanup and
 CC disposal of extra solvents and by-products is reduced. The fusion protein
 CC can also use directly different donor analogues and various acceptors
 CC with a terminal galactose residue

XX Sequence 828 B; 190 A; 241 C; 224 G; 173 T; 0 U; 0 Other;

Query Match 100.0%; Score 828; DB 2; Length 828;

Best Local Similarity 100.0%; Pred. No. 4,1e-242; Mismatches 0; Gaps 0;

Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAAAACCAAGCTTATGAGCTTACGCTTCGCGCAGAAACGACAGGCGACATTGCCGAT 60
 DB 1 ATGCAAAACCAAGCTTATGAGCTTACGCTTCGCGCAGAAACGACAGGCGACATTGCCGAT 60
 QY 61 ACCCTCGGAGGACAGCGGATCCCGTTTCAGTTTTCAGACCACTGATGCCGTCTGAAGG 120
 DB 61 ACCCTCGGAGGACAGCGGATCCCGTTTCAGTTTTCAGACCACTGATGCCGTCTGAAGG 120
 QY 121 CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTGTCCGCGACCCCTATTTCGACGGA 180
 DB 121 CTGGAACAGGCAATGCGGAACTCGTCCCGGCTTGTCCGCGACCCCTATTTCGACGGA 180
 QY 181 GTGGAAGAAAGCTCTTTATGAGCCAGCCGCTATTGTGGAAGAGGCAATTGACGAAGT 240
 DB 181 GTGGAAGAAAGCTCTTTATGAGCCAGCCGCTATTGTGGAAGAGGCAATTGACGAAGT 240
 QY 241 CTGCGCTATATACCGTATTTGAGAGCAGCTTTTATCTCGGGGAGAGGTAGGAAATTC 300
 DB 241 CTGCGCTATATACCGTATTTGAGAGCAGCTTTTATCTCGGGGAGAGGTAGGAAATTC 300
 QY 301 CTTCGCGAAGACGCTTGTGAGAAACGCTTTGACCCGGAATACCGCTTTATCGTCCGC 360
 DB 301 CTTCGCGAAGACGCTTGTGAGAAACGCTTTGACCCGGAATACCGCTTTATCGTCCGC 360
 QY 361 TTGGAAGAGATGTTATGACAGCTTCGACCTCGGCTCGGGTGGCGGATTAATCTCGGG 420
 DB 361 TTGGAAGAGATGTTATGACAGCTTCGACCTCGGCTCGGGTGGCGGATTAATCTCGGG 420
 QY 421 CGCGCTTTCGCGCTGTGGAAGACGAACTGCGGGGACGCGGGCTATATCTATTCGCCGA 480
 DB 421 CGCGCTTTCGCGCTGTGGAAGACGAACTGCGGGGACGCGGGCTATATCTATTCGCCGA 480
 QY 481 AAAGAGATCGGTTTTTCTGACAGGTTTTCGCGCTTCGCGCGGAGAGGCTGACCCG 540
 DB 481 AAAGAGATCGGTTTTTCTGACAGGTTTTCGCGCTTCGCGCGGAGAGGCTGACCCG 540
 QY 541 GTCGATCTGATGATGTTTCAGACATTTTTCGACAGGAGAAAGATGCCGATTTGCCAGCTC 600
 DB 541 GTCGATCTGATGATGTTTCAGACATTTTTCGACAGGAGAAAGATGCCGATTTGCCAGCTC 600
 QY 601 AATCCCGCTTGTGCGCCCAAGAGCTGCAATTAAGCAAGTTTCACGACCAAAACAGCGCA 660
 DB 601 AATCCCGCTTGTGCGCCCAAGAGCTGCAATTAAGCAAGTTTCACGACCAAAACAGCGCA 660
 QY 661 TTGGGAGCTGATGGAACAGACCGGCTTCGGAACCGGAAACAGAAAGGCGCATTC 720
 DB 661 TTGGGAGCTGATGGAACAGACCGGCTTCGGAACCGGAAACAGAAAGGCGCATTC 720
 QY 721 CCAGCAACATTCATTAACACCGGCTGATCCGCGCTTGAACCAAAATCAGACGGAAGG 780
 DB 721 CCAGCAACATTCATTAACACCGGCTGATCCGCGCTTGAACCAAAATCAGACGGAAGG 780

QY 781 GAAAGACCGCGCAAGGCGGACGATTCATTGTCCTTCCATTA 828
 DB 781 GAAAGACCGCGCAAGGCGGACGATTCATTGTCCTTCCATTA 828

RESULT 2

AAA81456/c
 ID AAA81456 strand; DNA; 46593 BP.

AC AAA81456;

DT 04-DEC-2000 (first entry)

DE N. meningitidis partial DNA sequence gnm_4 SEQ ID NO:4.

KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;

KW antigen; vaccine; diagnosis; infection; antibacterial; identification;

OS Neisseria meningitidis.

PN WO200022430-A2.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-US023573.

PR 09-OCT-1998; 98US-0103794P.

PR 30-APR-1999; 99US-0132068P.

PA (CHIR) CHIRON CORP.

PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC,

PI Masignani V, Galeotti C, Mora M, Ratti G, Scarcelll M, Scarlato V;

PI Rapunoli R, Pizza M;

DR MPI; 2000-318079/27.

PT Isolated nucleotide sequences of Neisseria meningitidis which can be used

PT in the diagnosis and treatment of N. meningitidis infection and other

PT Neisseria infections, for example, N.gonorrhoea.

PS Claim 7; Page 260-274; 1760dp; English.

CC The present invention describes methods of obtaining immunogenic proteins
 CC from Neisseria genomic sequences. AAA81453 to AAA82414 represent
 CC specifically claimed Neisseria meningitidis genomic DNA sequences;
 CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA
 CC sequences and their corresponding proteins; AAA81254 to AAA81259 and
 CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
 CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent
 CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
 CC used in the exemplification of the present invention. The nucleic acid
 CC sequences, protein sequences, and antibodies against them, can be used in
 CC the manufacture of a composition. The composition can be used as a
 CC medicament (or in the manufacture of a medicament) for treating,
 CC preventing or diagnosing infection due to Neisseria bacteria. For
 CC example, some of the identified proteins could be components of vaccines
 CC against Meningococcus B; against all serotypes; and/or against all
 CC pathogenic Neisseriae. Identification of sequences from the bacterium
 CC will also facilitate production of biological probes, particularly
 CC organism-specific probes. Attempts to make efficacious Meningococcus B
 CC vaccines have also been tried but none have successfully overcome
 CC antigenic variability. The provision of further, complete sequences may
 CC provide an opportunity to identify secreted or surface exposed proteins
 CC that may be presumed targets for the immune system and which are not
 CC antigenically variable or at least more conserved than other more
 CC variable regions

SEQ Sequence 46593 BP; 11355 A; 13195 C; 11355 G; 10687 T; 0 U; 1 Other;

Query Match 99.8%; Score 826.4; DB 3; Length 46593;

Mon Apr 10 08:45:27 2006

us-09-211-691-1.rng

Page 3

Best Local Similarity 99.9%; Pred. No. 7.1e-241;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGCAAAACACGTTATGAGCTTACGTTACGTTCCGCGCAGAAACGCGAGCGCATTTGCCGAT 60
Db 33536 ATGCAAAACACGTTATGAGCTTACGTTACGTTCCGCGCAGAAACGCGAGCGCATTTGCCGAT 33477
QY 61 ACCTTGGCAGGACGCGCATCCGCTTTCAGTTTTCGACGACATGATGCCGTCGAAAG 120
Db 33476 ACCTTGGCAGGACGCGCATCCGCTTTCAGTTTTCGACGACATGATGCCGTCGAAAG 33417
QY 121 CTGGAACAGGCAATGGCGAACTGTCGCCGCTTGTGCGGCAACCCCTATTTGAAGCGGA 180
Db 33416 CTGGAACAGGCAATGGCGAACTGTCGCCGCTTGTGCGGCAACCCCTATTTGAAGCGGA 33357
QY 181 GTGGAAGAAAGCCCTGTTATGAGCCACGCGTATGTGGAAGCAGAGCATTTGACGAAAGT 240
Db 33356 GTGGAAGAAAGCCCTGTTATGAGCCACGCGTATGTGGAAGCAGAGCATTTGACGAAAGT 33297
QY 241 CTGCGGTATATCAACCGTATTTGAGAGCAGCGTTTTCAGCGGCAAGGTGAGAAATTC 300
Db 33296 CTGCGGTATATCAACCGTATTTGAGAGCAGCGTTTTCAGCGGCAAGGTGAGAAATTC 33237
QY 301 CTGCGCAAGACGCTTGGCTGCAAGAACGCTTTGACCCGAGTACCGCTTTATGTCGCG 360
Db 33236 CTGCGCAAGACGCTTGGCTGCAAGAACGCTTTGACCCGAGTACCGCTTTATGTCGCG 33177
QY 361 TTGGAACAGATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGCGGATTAATCGGAG 420
Db 33176 TTGGAACAGATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGCGGATTAATCGGAG 33117
QY 421 CGGCGCTTTCGCGTGTGGAAGAGCACTGGGGGACGCGGGGCTATATCATTTCCCA 480
Db 33116 CGGCGCTTTCGCGTGTGGAAGAGCACTGGGGGACGCGGGGCTATATCATTTCCCA 33057
QY 481 AAAGCATGCGGTTTCTCTGACAGGTTTGCAGCCCTGCGCCGCAAGGCGTGCAACCC 540
Db 33056 AAAGCATGCGGTTTCTCTGACAGGTTTGCAGCCCTGCGCCGCAAGGCGTGCAACCC 32997
QY 541 GTGCAATCTGATGATGTTTTCAGCGCATTTTTCGACAGGAGAAATGCCGTTTCCAGCTC 600
Db 32996 GTGCAATCTGATGATGTTTTCAGCGCATTTTTCGACAGGAGAAATGCCGTTTCCAGCTC 32937
QY 601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCCAAGTTTTCAGCAACCAACAGCGCA 660
Db 32936 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCCAAGTTTTCAGCAACCAACAGCGCA 32877
QY 661 TTGGGAGCCTGATGCAACAGACGCGCTCTGAAACGCAAAAGGCGCGATTC 720
Db 32876 TTGGGAGCCTGATGCAACAGACGCGCTCTGAAACGCAAAAGGCGCGATTC 32817
QY 721 CCCGCAACACATTTCAACACGCGCTGATCGCGCTTTCGCAAAATCAGAGGAGAAAG 780
Db 32816 CCCGCAACACATTTCAACACGCGCTGATCGCGCTTTCGCAAAATCAGAGGAGAAAG 32757
QY 781 GAAAAACGCGGCAAGGCGCGCAACGTTTATGTCGCTTTCCAAATTA 828
Db 32756 GAAAAACGCGGCAAGGCGCGCAACGTTTATGTCGCTTTCCAAATTA 32709
```

RESULT 3

Continuation (6 of 9) of AAA81489 from base 500001 (N. meningitidis partial DNA sequence
WP Sequence split into 9 fragments LOCUS AAA81489 Accession Aaa81489

Fragment Name	Begin	End
AAA81489_0	1	110000
AAA81489_1	100001	210000
AAA81489_2	200001	310000
AAA81489_3	300001	410000
AAA81489_4	400001	510000
AAA81489_5	500001	610000
AAA81489_6	600001	710000
AAA81489_7	700001	810000

WP AAA81489_8 800001 837096
Query Match 99.9%; Score 826.4; DB 3; Length 110000;
Best Local Similarity 99.9%; Pred. No. 1e-240;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGCAAAACACGTTATGAGCTTACGTTACGTTCCGCGCAGAAACGCGAGCGCATTTGCCGAT 60
Db 91427 ATGCAAAACACGTTATGAGCTTACGTTACGTTCCGCGCAGAAACGCGAGCGCATTTGCCGAT 91368
QY 61 ACCTTGGCAGGACGCGCATCCGCTTTCAGTTTTCGACGACATGATGCCGTCGAAAG 120
Db 91367 ACCTTGGCAGGACGCGCATCCGCTTTCAGTTTTCGACGACATGATGCCGTCGAAAG 91308
QY 121 CTGGAACAGGCAATGGCGAACTGTCGCCGCTTGTGCGGCAACCCCTATTTGAAGCGGA 180
Db 91307 CTGGAACAGGCAATGGCGAACTGTCGCCGCTTGTGCGGCAACCCCTATTTGAAGCGGA 91248
QY 181 GTGGAAGAAAGCCCTGTTATGAGCCACGCGTATGTGGAAGCAGAGCATTTGACGAAAGT 240
Db 91247 GTGGAAGAAAGCCCTGTTATGAGCCACGCGTATGTGGAAGCAGAGCATTTGACGAAAGT 91188
QY 241 CTGCGGTATATCAACCGTATTTGAGAGCAGCGTTTTCAGCGGCAAGGTGAGAAATTC 300
Db 91187 CTGCGGTATATCAACCGTATTTGAGAGCAGCGTTTTCAGCGGCAAGGTGAGAAATTC 91128
QY 301 CTGCGCAAGACGCTTGGCTGCAAGAACGCTTTGACCCGAGTACCGCTTTATGTCGCG 360
Db 91127 CTGCGCAAGACGCTTGGCTGCAAGAACGCTTTGACCCGAGTACCGCTTTATGTCGCG 91068
QY 361 TTGGAACAGATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGCGGATTAATCGGAG 420
Db 91067 TTGGAACAGATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGCGGATTAATCGGAG 91008
QY 421 CGGCGCTTTCGCGTGTGGAAGAGCACTGGGGGACGCGGGGCTATATCATTTCCCA 480
Db 91007 CGGCGCTTTCGCGTGTGGAAGAGCACTGGGGGACGCGGGGCTATATCATTTCCCA 90948
QY 481 AAAGCATGCGGTTTCTCTGACAGGTTTGCAGCCCTGCGCCGCAAGGCGTGCAACCC 540
Db 90947 AAAGCATGCGGTTTCTCTGACAGGTTTGCAGCCCTGCGCCGCAAGGCGTGCAACCC 90888
QY 541 GTGCAATCTGATGATGTTTTCAGCGCATTTTTCGACAGGAGAAATGCCGTTTCCAGCTC 600
Db 90887 GTGCAATCTGATGATGTTTTCAGCGCATTTTTCGACAGGAGAAATGCCGTTTCCAGCTC 90828
QY 601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCCAAGTTTTCAGCAACCAACAGCGCA 660
Db 90827 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCCAAGTTTTCAGCAACCAACAGCGCA 90768
QY 661 TTGGGAGCCTGATGCAACAGACGCGCTCTGAAACGCAAAAGGCGCGATTC 720
Db 90767 TTGGGAGCCTGATGCAACAGACGCGCTCTGAAACGCAAAAGGCGCGATTC 90708
QY 721 CCCGCAACACATTTCAACACGCGCTGATCGCGCTTTCGCAAAATCAGAGGAGAAAG 780
Db 90707 CCCGCAACACATTTCAACACGCGCTGATCGCGCTTTCGCAAAATCAGAGGAGAAAG 90648
QY 781 GAAAAACGCGGCAAGGCGCGCAACGTTTATGTCGCTTTCCAAATTA 828
Db 90647 GAAAAACGCGGCAAGGCGCGCAACGTTTATGTCGCTTTCCAAATTA 90600
```

RESULT 4

AAF21612/c
ID AAF21612 standard; DNA; 349980 BP.

Fragment Name	Begin	End
AAF21612;	1	349980
13-MAR-2001 (first entry)	1	349980
Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.	1	349980

KM *Neisseria meningitidis*; *Neisseria gonorrhoeae*; immunogenic; vaccine;
 KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
 KW ds.
 OS *Neisseria meningitidis*.
 XX MO200066791-A1.
 XX
 PD 09-NOV-2000.
 XX
 PF 08-MAR-2000; 2000MO-US005928.
 XX
 PR 30-APR-1999; 99US-0132068P.
 PR 08-OCT-1999; 99MO-US023573.
 XX 28-FEB-2000; 2000GB-00004693.
 XX
 PA (CHIR) CHIRON CORP.
 XX (GENO-) INST GENOMIC RES.
 XX
 PI Pizza M, Hickey E, Peterson J, Tetrelin H, Venter JC;
 PI Masignani V, Galeotti C, Mora M, Ratti G, Scarcelll M, Scarlato V;
 PI Rapuoli R, Frazer CM, Grandi G;
 XX
 DR WPI, 2000-6476q3/62.
 XX
 PT *Neisseria meningitidis* B full length genome sequence and open reading
 PT frames are used to detect, treat and prevent *Neisseria* infections.
 XX
 PS Claim 7; Appendix A; 692pp; English.
 XX
 CC The present invention describes the full length genome of *Neisseria*
 CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
 CC represent fragments of the NMB genomic sequence, as the sequence was too
 CC long to go in a record on its own it was split into 8 sequences which
 CC overlap each other at the beginning and end of each sequence by 49980 bp
 CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of
 CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
 CC AAF21608, and so on). AAF21545 to AAF21588 encode the *Neisseria* proteins
 CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
 CC primers which are used in the exemplification of the present invention.
 CC The NMB genome and fragments from it have antibacterial activity, and can
 CC be used in vaccines and gene therapy. *Neisseria* nucleic acids, proteins
 CC and/or antibodies which binds to the proteins can be used in compositions
 CC for treating or preventing infection due to *Neisseria* bacteria or as a
 CC diagnostic reagent for detecting the presence of *Neisseria* bacteria or
 CC of antibodies raised to *Neisseria* bacteria. Computers, computer memory,
 CC computer storage medium or computer databases can be used in a search to
 CC identify open reading frames (ORFs) or coding sequences within the NMB
 CC genome. The DNA sequences provide further opportunities to find antigenic
 CC or immunogenic proteins which are more effective in vaccines than the
 CC outer membrane proteins currently used
 CC
 XX
 SQ Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 U; 0 Other;
 XX
 Query Match 99.8%; Score 826.4; DB 3; Length 349980;
 Best Local Similarity 99.9%; Pred. No. 1.7e-240;
 Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 ATGCAAAACCACTTATGAGCTTACGTTAGCTTCCGCGAAGACGAGCGGCAATTGCCGAT 60
 DB 226656 ATGCAAAACCACTTATGAGCTTACGTTAGCTTCCGCGAAGACGAGCGGCAATTGCCGAT 226597
 QY 61 ACCCTGCGAGGCAAGGCAATCCGCTTTCAGTTTTCAGTTCGACGACGTAATGCCGTAAGG 120
 DB 226596 ACCCTGCGAGGCAAGGCAATCCGCTTTCAGTTTTCGACGACGTAATGCCGTAAGG 226537
 QY 121 CTGGAACAGGCAATGCGGCAACTGTCCTCCGCGCTTGTGCGCGCACCCCTATTGAGCGGA 180
 DB 226536 CTGGAACAGGCAATGCGGCAACTGTCCTCCGCGCTTGTGCGCGCACCCCTATTGAGCGGA 226477
 QY 181 GTGGAAAAAGCCTGCTTATGAGCAAGCGCGGATTTGTGAAGCAGGCAATGGAGAAGT 240
 DB 226476 GTGGAAAAAGCCTGCTTATGAGCAAGCGCGGATTTGTGAAGCAGGCAATGGAGAAGT 226417

QY 241 CTGCCGTATATACCGGTATTTGAGGACGACGTTTACTCGGCGAAGGTGAGAAAAATTC 300
 DB 226416 CTGCCGTATATACCGGTATTTGAGGACGACGTTTACTCGGCGAAGGTGAGAAAAATTC 226557
 QY 301 CTTCGCGAAGACGCTTGCGCTGCAGACGCTTTGACCCCGATACCCTTTATCGTCCGC 360
 DB 226356 CTTCGCGAAGACGCTTGCGCTGCAGACGCTTTGACCCCGATACCCTTTATCGTCCGC 226297
 QY 361 TTGAAAGCATGTTATGACAGTCTGACCTCGGCTCGGCGGAGGAGATTAATGACGGG 420
 DB 226296 TTGAAAGCATGTTATGACAGTCTGACCTCGGCTCGGCGGAGGAGATTAATGACGGG 226237
 QY 421 CGCGCTTTCCGCTGTTGAAAGCGAACACTGGGGGACGCGGCGCTATATGATTTCCCGA 480
 DB 226236 CGCGCTTTCCGCTGTTGAAAGCGAACACTGGGGGACGCGGCGCTATATGATTTCCCGA 226177
 QY 481 AAAGCATGCGGTTTTCCTGACAGGTTTTCGCGCTTCGCGCGGAGGGCTGACCCG 540
 DB 226176 AAAGCATGCGGTTTTCCTGACAGGTTTTCGCGCTTCGCGCGGAGGGCTGACCCG 226117
 QY 541 GTGATCTGATGATGTTGACGATTTTTCGACAGGAGGAGATGCGTTTGCAGCTC 600
 DB 226116 GTGATCTGATGATGTTGACGATTTTTCGACAGGAGGAGATGCGTTTGCAGCTC 226057
 QY 601 AATCCGCGCTTGCGCGCCCAAGAGCTGATATGCAAGTTTTCACGACCAAAAGCGCA 660
 DB 226056 AATCCGCGCTTGCGCGCCCAAGAGCTGATATGCAAGTTTTCACGACCAAAAGCGCA 225997
 QY 661 TTGGCAGCTGATGACACGACGCGCTCTCTGAAACGCAAAAGCGCGATTC 720
 DB 225996 TTGGCAGCTGATGACACGACGCGCTCTCTGAAACGCAAAAGCGCGATTC 225937
 QY 721 CCCGCAACATTTCAACACCGCTGATCCGCGCTTGACCAAAATGACGAGGAAAG 780
 DB 225936 CCCGCAACATTTCAACACCGCTGATCCGCGCTTGACCAAAATGACGAGGAAAG 225877
 QY 781 GAAAAACGCGCGAAGCGGACAGTTCATTTGCTTCCAAATA 828
 DB 225876 GAAAAACGCGCGAAGCGGACAGTTCATTTGCTTCCAAATA 225829

RESULT 5
 AAT14061
 ID AAT14061 standard; DNA; 5859 BP.
 XX
 AC AAT14061;
 XX
 DT 16-OCT-2003 (revised)
 DT 09-JUL-1996 (first entry)
 XX
 XX N. gonorrhoeae 1gt locus.
 XX
 XX
 XX Glycosyltransferase; 1,4p-oligosaccharide; 1gt gene; LOS locus; vaccine;
 KW ds.
 XX
 OS *Neisseria gonorrhoeae*, strain F62.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 445..1491
 FT /tag= a
 FT /product= "LgtA"
 FT 699..715
 FT /tag= b
 FT /note= "poly-G tract"
 FT 1491..2330
 FT /tag= c
 FT /product= "LgtB"
 FT 2342..3262
 FT /tag= d
 FT /product= "LgtC"
 FT 2499..2508
 FT misc_feature
 FT /tag= e

```

FT      /note= "poly-G tract"
FT      3322..4335
FT      /*tag= f
FT      /product= "lgtD"
FT      misc_feature
FT      3576..3586
FT      /*tag= g
FT      /note= "poly-G tract"
FT      CDS
FT      4354..5196
FT      /*tag= h
FT      /product= "lgtE"
XX
XX      MO9610086-A1.
XX
XX      04-APR-1996.
XX
XX      25-SEP-1995; 95WO-US012317.
XX
XX      26-SEP-1994; 94US-00312387.
XX
XX      (UYRQ ) UNIV ROCKEFELLER.
XX
XX      Gotschlich EC;
XX
XX      MPI; 1996-200924/20.
XX      P-PSDB; AAR91311, AAR91312, AAR91313, AAR91314, AAR91315.
XX
XX      Nucleic acids encoding glycosyl transferase(s) - used in the diagnosis of
XX      infection with Neisseria and for the biosynthesis of oligo:saccharide(s).
XX      claim 1; Fig 29-m; 81pp; English.
XX
XX      The lgt locus (AAT91061) of Neisseria gonorrhoeae F62 contains 5 open
XX      reading frames, lgtA, lgtB, lgtC, lgtD and lgtE, coding for 5
XX      glycosyltransferases (see also AAR91311-15) involved in gonococcal
XX      lipooligosaccharide (LOS) biosynthesis. The sequence was constructed from
XX      2 clones isolated from an F62 gene bank in lambda-200 following screening
XX      with plasmid pRI0P1. 3 Of the coding sequences contain poly-G tracts that
XX      make them susceptible to premature termination. The lgt coding sequences
XX      CC can be used for prodn. of recombinant Lgt glycosyltransferases that are
XX      utilised in biosynthesis of LOS useful in vaccine prepn. (Updated on 16-
XX      OCT-2003 to standardise OS field)
XX
XX      Sequence 5859 BP; 1411 A; 1462 C; 1661 G; 1325 T; 0 U; 0 Other;
XX
XX      Query Match      87.7%; Score 726; DB 2; Length 5859;
XX      Best Local Similarity 93.2%; Pred. No. 1,4e-210;
XX      Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
XX
XX      1 AATCAAAACCAAGTATACAGCTTAGCTCCGCGCAGAAAGGCGGCGCATTTGCGCAT 60
XX      1491 ATGCAAAACCAAGTATACAGCTTAGCTCCGCGCAGAAAGGCGGCGCATTTGCGCAT 1550
XX
XX      61 ACCTTGCGGAGGCGGCGCATCCCGTTTCAGTTTTCAGCGACATGATGCGTCTGAAGG 120
XX      1551 ACCTTGCGGAGGCGGCGCATCCCGTTTCAGTTTTCAGCGACATGATGCGTCTGAAGG 1610
XX
XX      121 CTGGAACAGGCAATGCGGCACTGTCGCCGCTTGTCGCGCACCCCTATTTGAGCGGA 180
XX      1611 CTGGAACAGGCAATGCGGCACTGTCGCCGCTTGTCGCGCACCCCTATTTGAGCGGA 1670
XX
XX      181 GTGGAAGAAAGCGCTTTATGAGCGAGCGGTATTTGGAAGAGGAGCATTTGAGCGAAGGT 240
XX      1671 GTGGAAGAAAGCGCTTTATGAGCGAGCGGTATTTGGAAGAGGAGCATTTGAGCGAAGGT 1730
XX
XX      241 CTGCGGTATATCAACGCTATTTGAGAGCAAGTTTTCAGCGCAAGGTGAGGAAATTC 300
XX      1731 GTACCGTATATCGCCGATTTTGAAGATGATCTTACTGCGGCAAGGCGGAGGAGCTTC 1790
XX
XX      301 CTGCGCGAAGAGCGCTTGCTGCAAGAGCGTTTGAACCGGATACCGCTTTATGTCGCCG 360
XX      1791 CTGCGCGAAGAGATCTTGCTGCAAGAGCGTTTGAACCGGATACCGCTTTATGTCGCCG 1850
XX
XX      361 TTGGAAGAGATGTTATGACAGTCTGACTCGCCCTCGGCGGTGCGGATTAATCGCGGG 420

```

```

Db      1851 TTGGAAGAGATGTTATGACAGTCTGACTCGCCCTCCGGGCTGCGGACTACGCGGG 1910
XX
XX      421 CGGCGCTTTCCGCTGTGGAAGAGCAACACTGCGGAGGAGCGCGGCTATATCATTTCCCGA 480
XX      1911 CGGCGCTTTCCGCTGTGGAAGAGCAACACTGCGGAGGAGCGCGGCTATATCATTTCCCGA 1970
XX
XX      481 AAGGAGAGCGGTTTTTCTGACAGGTTTGCGCGCTTCGCGCGGAGGAGGCTGACCCG 540
XX      1971 AAGGAGAGCGGTTTTTCTGACAGGTTTGCGCGCTTCGCGCGGAGGAGGCTGACCCG 2030
XX
XX      541 GTGCAATCTGATGATGTTTTCAGCAATTTTTTCAGACAGGAGAGATGCGGTTTCCAGCTC 600
XX      2031 GTGCAATCTGATGATGTTTTCAGCAATTTTTTCAGACAGGAGAGATGCGGTTTCCAGCTC 2090
XX
XX      601 AATCCCGCTTGTGCGCGCCCAAGAGCTGCAATTAATGCCAAGTTTTCAGCAACCAACAGCGCA 660
XX      2091 AATCCCGCTTGTGCGCGCCCAAGAGCTGCAATTAATGCCAAGTTTTCAGCAACCAACAGCGCA 2150
XX
XX      661 TTGGGAGAGCTGATGCAACAGACGCGCTTCTGAAACCGCAACAGCAAGGCGCATTTCC 720
XX      2151 TTGGGAGAGCTGATGCAACAGACGCGCTTCTGAAACCGCAACAGCAAGGCGCATTTCC 2210
XX
XX      721 CCGGCAACACATTCACACACCGCGCTGATCCGCGCTTTCAGCAACCAATTCAGAGGAGG 780
XX      2211 CCGGCAACACATTCACACACCGCGCTGATCCGCGCTTTCAGCAACCAATTCAGAGGAGG 2270
XX
XX      781 GAAAAACGCGCGCAAGGCGGCAACAGTTTCTG 814
XX      2271 GAAAAACGCGCGCAAGGCGGCAACAGTTTCTG 2304

```

```

XX      RESULT 6
XX      AAT9230
XX      ID AAT9230 standard; DNA; 5859 BP.
XX
XX      AAT9230;
XX
XX      AC 17-OCT-2003 (revised)
XX      DT 21-MAR-1997 (first entry)
XX
XX      DE Lipo-oligosaccharide (including polyglycosyltransferase) gene.
XX
XX      KM Polyglycosyltransferase; N-acetylglucosaminyl transferase;
XX      KM N-acetylglucosaminyl transferase; lipo-oligosaccharide; ss.
XX
XX      OS Neisseria gonorrhoeae; ATCC 33084.
XX
XX      FH Key      Location/Qualifiers
XX      FT CDS      1..381
XX      FT          /*tag= a
XX      FT CDS      445..1491
XX      FT          /*tag= b
XX      FT          /product= "polyglycosyltransferase"
XX      FT CDS      2342..3262
XX      FT          /*tag= c
XX      FT CDS      3322..4335
XX      FT          /*tag= d
XX      FT CDS      4354..5196
XX      FT          /*tag= e
XX
XX      PN MO9640971-A1.
XX      PD 19-DEC-1996.
XX
XX      PF 03-JUN-1996; 96WO-US008323.
XX
XX      PR 07-JUN-1995; 95US-00478140.
XX
XX      PA (NEOS-) NROSE TECHNOLOGIES INC.
XX      PI Johnson KF, Roth S, Buczala SL;

```

DR WPI, 1997-052351/05.
 P-PSDB; AAM06576, AAM06577, AAM06578, AAM06579, AAM06580.
 XX Transfer of at least 2 saccharide units using poly(glycosyl:transferase -
 PT isolated from *N. gonorrhoeae*, catalyses the addition of both GlcNAc and
 PT GalNAc di:saccharide(s) units to a single galactose moiety.
 PS
 XX Disclosure; Fig 2A-H; 38pp; English.

CC A lipooligosaccharide-encoding gene region (AAT9230) of *Neisseria*
 CC gonorrhoeae ATCC 35084 includes a coding sequence for a novel N-
 CC polyglycosyltransferase (PGTase) (AAM06576) that catalyses the addition
 CC of both GlcNAc and GalNAc disaccharides to a single galactose moiety. The
 CC products (AAM06577-80) of the other coding sequences of this region are
 CC not identified. The PGTase gene can be obtd. using standard techniques
 CC and incorporated into a vector to allow prodn. of the PGTase in
 CC transformed host cells. The enzyme is useful in the synthesis of
 CC oligosaccharides. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 5859 BP; 1407 A; 1462 C; 1661 G; 1329 T; 0 U; 0 Other;

Query Match 87.7%; Score 726; DB 2; Length 5859;
 Best Local Similarity 93.2%; Pred. No. 1.4e-210;
 Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGCAAAACACGCTTATGAGCTTACGCTTCCGCGCAAAACGAGGCGCAATTGCCGAT 60
 DB 1491 ATGCAAAACACGCTTATGAGCTTACGCTTCCGCGCAAAACGAGGCGCAATTGCCGCA 1550
 QY 61 ACCTTGCGAGGACGCGCATCCCGTTTCAGTTTTCAGACGACATGACGCTGTAAGG 120
 DB 1551 ACCTTGCGAGGATGCGCGCATCCCGTTTCAGTTTTCAGACGACATGACGCTGTAAGG 1610
 QY 121 CTGGAACAGGCAATGCGGAACTCGCTCCGCTTCCGCGCAACCCCTATTGAGCGGA 180
 DB 1611 CTGGAACGCGCAATGCGGAACTCGCTCCGCTTCCGCGCAACCCCTATTGAGCGGA 1670
 QY 181 GTGGAAGAAAGCTGCTTTATGAGCCAGCGCTTATGAGAGCGATTTGACGAAAGT 240
 DB 1671 GTGGAAGAAAGCTGCTTTATGAGCCAGCGCTTATGAGAGCGATTTGACGAAAGG 1730
 QY 241 CTGCGCTATATCAACGCTTATTTGAGAGCGATTTTATCTGCGGAGAGGATGAAATTC 300
 DB 1731 GTAACCTATATGCGCTTATTTGAGAGATGATGCTTACTGCGGAGAGCGCGAGCAATTC 1790
 QY 301 CTTCGCGAAGAGCTTGGCTGCAAGAAAGCTTTGACCTGAGATACCGCTTTATGCTCCG 360
 DB 1791 CTTCGCGAAGATACCTTGGCTGCAAGAAAGCTTTGACCTGAGATACCGCTTTGCTCCG 1850
 QY 361 TTGGAACAGATGTTTATGACAGCTGCAAGCTGCGCTCCGCGGAGGCGGATTAATCGCGG 420
 DB 1851 TTGGAACAGATGTTTATGACAGCTGCAAGCTGCGCTCCGCGGAGGCGGATTAATCGCGG 1910
 QY 421 CGCGCTTTCCGCTGTTGGAAGCCGAACCTGCGGAGCGCGGCTATATCATTTCCCGA 480
 DB 1911 CGCGCTTTCCGCTTTTGGAAAGCCGAACCTGCGGAGCGCGGCTATATTTTCCCGA 1970
 QY 481 AAAGGATGCGGTTTTTCTGGAAGGTTTTCGCGCTGCGCGCGGAGAGGCTGACCC 540
 DB 1971 AAAGGATGCGGTTTTTCTGGAAGGTTTTCGCGCTGCGCGCGGAGAGGCTGACCC 2030
 QY 541 GTGCAATCGATGATGTTGACGCAATTTTTCGACAGGAAAGATGCGGTTTCCAGCTC 600
 DB 2031 GTGCAATGATGATGTTGCGCAACCTTACGACAGGAAAGATGCGGTTTCCAGCTC 2090
 QY 601 AATCCCGCTTGGCGCCCAAGAGCTGATATGCAAGTTTCAAGACCAAAACAGCGCA 660
 DB 2091 AATCCCGCTTGGCGCCCAAGAGCTGATATGCAAGTTTCAAGACCAAAACAGCGCA 2150
 QY 661 TTGGGACGCTGATGAAACAGACGCGCTCTGAAACCGCAACAGCAAAAGCGGATTC 720
 DB 2151 TTGGGACGCTGATGAAACAGACGCGCTCTGAAACCGCAACAGCAAAAGCGGATTC 2210

QY 721 CCCGCAACACATTCAAACACCGCTGATCCGCGCTTGACCAAAATGACGAGGAAAG 780
 DB 2211 CCCGCAACACATTCAAACACCGCTGATCCGCGCTTGACCAAAATGACGAGGAAAG 2270
 QY 781 GAAAAACCGCGCAAAAGCGCGGAGACATGTTCAATTG 814
 DB 2271 GAAAAACCGCGCAAAAGCGCGGAGACATGTTATCG 2304

RESULT 7
 ABZ38360
 ID ABZ38360 standard; DNA; 837 BP.
 XX
 AC ABZ38360;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE N. gonorrhoeae nucleotide sequence SEQ ID 1309.
 XX
 KW Antibacterial; infection; vaccine; gene therapy; gene; de.
 XX
 OS *Neisseria gonorrhoeae*.
 XX
 PN W0200279243-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 12-FEB-2002; 2002WO-1B002069.
 XX
 PR 12-FEB-2001; 2001GB-00003424.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Fontana MR, Pizza M, Maignani V, Monaci E;
 DR WPI; 2003-058415/05.
 XX
 DR P-PSDB; ABP77390.

XX New protein from *Neisseria gonorrhoeae*, useful for the manufacture of a
 PT medicament for treating or preventing *N. gonorrhoeae* infection.
 XX
 PS Disclosure; Page 283; 815pp; English.
 CC
 CC The present invention relates to proteins from *Neisseria gonorrhoeae*.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing *N. gonorrhoeae*
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records AB237706-AB242016 represent nucleic acid
 CC molecules of the invention
 XX

XX Sequence 837 BP; 186 A; 236 C; 233 G; 182 T; 0 U; 0 Other;

Query Match 87.1%; Score 721.2; DB 10; Length 837;
 Best Local Similarity 92.9%; Pred. No. 1.9e-209;
 Matches 756; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 ATGCAAAACACGCTTATGAGCTTACGCTTCCGCGCAAAACGAGGCGCAATTGCCGAT 60
 DB 1 ATGCAAAACACGCTTATGAGCTTACGCTTCCGCGCAAAACGAGGCGCAATTGCCGCA 60
 QY 61 ACCTTGCGAGGACGCGCATCCCGTTTCAGTTTTCAGACGACATGACGCTGTAAGG 120
 DB 61 ACCTTGCGAGGATGCGCGCATCCCGTTTCAGTTTTCAGACGACATGACGCTGTAAGG 120
 QY 121 CTGGAACAGGCAATGCGGAACTCGCTCCGCTTGGCGCAACCCCTATTGAGCGGA 180
 DB 121 CTGGAACAGGCAATGCGGAACTCGCTCCGCTTGGCGCAACCCCTATTGAGCGGA 180
 QY 181 GTGGAAGAAAGCTGCTTTATGAGCCAGCGCTTATGAGAGCGGATTTGAGCAAGGT 240
 DB 181 GTGGAAGAAAGCTGCTTTATGAGCCAGCGCTTATGAGAGCGGATTTGAGCAAGGT 240


```
QY 241 CTGCGGTATATACCGTATTGAGACGAGCTTTTACTCGGCGAAGGTGAGAAATTC 300
DB 241 CTGCGGTATATACCGGTATTGAGACGAGCTTTTACTCGGCGAAGGTGAGAAATTC 300
QY 301 CTGCGCGAAGACGCTTGGCTGCAAGAACGCTTTGACCCGATACCGCTTTATCTGTCGC 360
DB 301 CTGCGCGAAGATCTTGGTTGGAAGAGCTTTGATTAAGATTCGCGCTTTATCTGTCGC 360
QY 361 TTGGAAACGATGTTTATGACGTCCTGACCTCGCCCTCGCGCTGCGGATTACTGCGGG 420
DB 361 TTGGAAACGATGTTTATGACGTCCTGACCTCGCCCTCGCGCTGCGGATTACTGCGGG 420
QY 421 CGGCGCTTTCGCGCTGTTGAAAGGAAACACTGCGGAGGAGCGGCTATATATTCCTCGA 480
DB 421 CGGCGCTTTCGCGCTTTCGAAAGGAAACACTGCGGAGGAGCGGCTATATATTCCTCGA 480
QY 481 AAGCGATGCGGCTTTTCTGCAAGGTTTCCGCGCTCGCGCGGAGGCTGCAACCC 540
DB 481 AAGCGATGCGGCTTTTCTGCAAGGTTTCCGCGCTCGCGCGGAGGCTGCAACCC 540
QY 541 GTGATCTGATGATGTTTCAAGCAATTTTTCAGAGGAGAAATGCGGTTTCCAGCTC 600
DB 541 GTGATCTGATGATGTTTCAAGCAATTTTTCAGAGGAGAAATGCGGTTTCCAGCTC 600
QY 601 AATCCGCGCTTTCGCGCGGAGGCTGCAATTAAGCAAGTTTTCAGACCAAAACAGCGCA 660
DB 601 AATCCGCGCTTTCGCGCGGAGGCTGCAATTAAGCAAGTTTTCAGACCAAAACAGCGCA 660
QY 661 TTGGGAGCGCTGATTCGACAGCAGCGCTCTCTGTAACCGCAAGAGGCGCGATTC 720
DB 661 TTGGGAGCGCTGATTCGACAGCAGCGCGCTCTCTGTAACCGCAAGAGGCGCGATTC 720
QY 721 CCGCGCAACATTCGAAACACCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAG 780
DB 721 CCGCGCAACATTCGAAACACCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAG 780
QY 781 GAAAAACGCGCGCAAGGCGCGAAGCAATTCG 814
DB 781 GAAAAACGCGCGCAAGGCGCGAAGCAATTCG 814

RESULT 8
ABZ38368
ID ABZ38368 standard; DNA; 840 BP.
XX
AC ABZ38368;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae nucleotide sequence SEQ ID 1325.
XX
KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX
OS Neisseria gonorrhoeae.
XX
PN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-1B002069.
XX
PR 12-FEB-2001; 2001GB-00003424.
XX
PI (CHIR-) CHIRON SPA.
XX
PA Fontana MR, Pizsa M, Masignani V, Monaci E;
XX
DR MPI; 2003-058415/05.
XX
DR P-PSDB; ABP77398.
XX
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
medicament for treating or preventing N. gonorrhoeae infection.
```

```
XX
PS Disclosure; Page 284; 815bp; English.
XX
CC The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
CC molecules of the invention
XX
SQ Sequence 840 BP; 209 A; 167 C; 242 G; 222 T; 0 U; 0 Other;
Query Match 57.0%; Score 471.8; DB 10; Length 840;
Best Local Similarity 80.7%; Pred. No. 2.7e-133;
Matches 551; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY 1 ATGCAAAACGAGTATGAGCTTAGCTTCGCGGAGAAAGGAGGCGCAATTCGCGAT 60
DB 1 ATGCAAAACGAGTATGAGCTTAGCTTCGCGGAGAAAGGAGGCGCAATTCGCGAT 60
QY 61 ACCCTCGGAGGAGCGGATCCGTTTCAAGTTTTCAGCACTGATGCGCTGCAAG 120
DB 61 ACCCTCGGAGGAGCGGATCCGTTTCAAGTTTTCAGCACTGATGCGCTGCAAG 120
QY 121 CTGAAACAGGCAATGCGGAACTGCTCCGCGCTTGTGCGGCAACCCCTATTGACGGA 180
DB 121 CTGAAACAGGCAATGCGGAACTGCTCCGCGCTTGTGCGGCAACCCCTATTGACGGA 180
QY 181 GTGAAAAACGCTGCTTATGAGCAAGCGCGATTAATGGAACAGCAATTCGAAAGT 240
DB 181 GTGAAAAACGCTGCTTATGAGCAAGCGCGATTAATGGAACAGCAATTCGAAAGT 240
QY 241 CTGCGGTATATACCGTATTGAGACGAGCTTTTACTCGGCGAAGGTGAGAAATTC 300
DB 241 CTGCGGTATATACCGTATTGAGACGAGCTTTTACTCGGCGAAGGTGAGAAATTC 300
QY 301 CTGCGCGAAGACGCTTGGCTGCAAGAGCTTTGACCCGATACCGCTTTATCGTCGC 360
DB 301 CTGCGCGAAGATCTGTTGGAAGAGCGCTTTGATTAAGATTCGCTTTATCGTCGC 360
QY 361 TTGAAAGAGATGTTATGACAGCTCCGACCTCGCGCTCGGCGGATTAATTCGCGG 420
DB 361 TTGAAAGAGATGTTATGACAGCTCCGACCTCGCGCTCGGCGGATTAATTCGCGG 420
QY 421 CGGCGCTTTCGCTGTTGAAAGCGAACACTGCGGAGCGCGGCTATATTCGCGA 480
DB 421 CGGCGCTTTCGCTGTTGAAAGCGAACACTGCGGAGCGCGGCTATATTCGCGA 480
QY 481 AAGCGATGCGGCTTTTCTGCAAGGTTTCCGCGCTCGCGCGGAGGCTGCAACCC 540
DB 481 AAGCGATGCGGCTTTTCTGCAAGGTTTCCGCGCTCGCGCGGAGGCTGCAACCC 540
QY 541 GTGATCTGATGATGTTTCAAGCAATTTTTCAGAGGAGAAATGCGGTTTCCAGCTC 600
DB 541 GTGATCTGATGATGTTTCAAGCAATTTTTCAGAGGAGAAATGCGGTTTCCAGCTC 600
QY 601 AATCCGCGCTTTCGCGCGGAGGCTGCAATTAAGCAAGTTTTCAGACCAAAACAGCGCA 660
DB 601 AATCCGCGCTTTCGCGCGGAGGCTGCAATTAAGCAAGTTTTCAGACCAAAACAGCGCA 660
QY 661 TTGGGAGCGCTGATTCGACAGCAGCGCTCTCTGTAACCGCAAGAGGCGCGATTC 720
DB 661 TTGGGAGCGCTGATTCGACAGCAGCGCTCTCTGTAACCGCAAGAGGCGCGATTC 720
QY 721 GAAAAACGCGCGCAAGGCGCGAAGCAATTCG 814
DB 721 GAAAAACGCGCGCAAGGCGCGAAGCAATTCG 814

RESULT 9
ABX09919
ID ABX09919 standard; DNA; 840 BP.
XX
AC ABX09919;
XX
```

27-OCT-2003 (revised)
22-JAN-2003 (first entry)
N. meningitidis DNA encoding a vaccine antigen #61.
Vaccine; antigen; ds; meningococcal disease; pathogenic bacteria;
meningitis.
Neisseria meningitidis serogroup B.
MO20027648-A2.
03-OCT-2002.
22-MAR-2002; 2002MO-GB001399.
22-MAR-2001; 2001GB-00007219.
(MICR-) MICROBIOLOGICAL RES AUTHORITY.
(IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
Robinson A, Goringe AR, Hudson MJ, Bracegirdle P, West DM;
Oliver KJ, Kroll JS, Langford PR;
WPI; 2003-018958/01.
P-PSDB; AB06084.
Identifying an antigen for manufacturing a vaccine against meningococcal
infection, comprises contacting antibodies with polypeptides, detecting
polypeptide-antibody complexes, and identifying bound polypeptides as
antigens.
Claim 34; Page 278-279; 310pp; English.
The invention relates to identifying an antigen comprising: (a) obtaining
antibodies against a commensal bacteria, or an extract from a commensal
bacteria; (b) contacting the antibodies with polypeptides obtained from
an expression library of either a commensal or a pathogenic bacteria; (c)
determining whether the polypeptides bind to antibodies; and (d) where a
polypeptide binds to an antibody) identifying that polypeptide as an
antigen. Also included are: (1) a method of preparing a vaccine
composition, comprising identifying an antigen with the above method, and
combining the antigen with a carrier; (2) a vaccine composition obtained
by the above methods; (3) an antigen identified by the above method; (4)
a polypeptide encoded by all or a part of a nucleic acid sequence
comprising the Neisseria lactamica DNA sequences detailed in the
specification; (5) a vector comprising the nucleic acid molecule; (6) a
method of preparing a composition for vaccination against infection by
pathogenic bacteria, comprising: (a) obtaining a first antigen from a
commensal Neisseria; (b) comparing the amino acid sequence of the first
antigen with the amino acid sequence of the second antigen from a
pathogenic bacteria, or comparing the sequence of a nucleic acid which
codes for the first antigen with the sequence of the nucleic acid that
codes for the second antigen; and if the first antigen is homologous to
the second antigen or if the nucleic acid sequence for the first antigen
is homologous to that of the second antigen, and (c) preparing a
composition for vaccination against bacterial infection comprising the
first antigen; and (7) an antibody that binds to the polypeptide antigen.
The method is useful in screening commensal and pathogenic bacteria for
previously unidentified vaccine antigens by identifying polypeptide
antigens that bind to sera raised against commensal bacterial proteins.
The polypeptide is useful as a vaccine antigen which may be used in the
manufacture of a medicament for vaccination against meningococcal
infection (e.g. meningitis). The present sequence encodes an antigenic
protein from the pathogenic bacteria N. meningitidis. (Updated on 27-OCT-
2003 to standardise OS field)

Sequence 840 BP; 218 A; 154 C; 244 G; 224 T; 0 U; 0 Other;

Query Match 51.4%; Score 426; DB 8; Length 840;
Best Local Similarity 75.2%; Pred. No. 2.7e-119;
Matches 531; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 1 ATCAAAACCAAGTATGAGTTCAGTTCCTCCGCGGAGAACGCGAGGCGCACATGCGCAT 60
DB 1 ATGCAAAACCAAGTATGAGTTCAGTTCCTCCGCGGAGAACGCGAGGCGCACATGCGCAT 60
QY 61 ACCCTGCGGAGGAGCGGAGTCCGTTTCAAGTTTTCAGCGCATGATCCGCTGAAAG 120
DB 61 ACCCTGCGGAGGAGGAGTCCGTTTCAAGTTTTCAGCGCATGATCCGCTGAAAG 120
QY 121 CTGGAACGAGCAATGCGGAGTCCGTTTTCAGCGCATGATCCGCTGAAAG 180
DB 121 CTGGAACGAGCAATGCGGAGTCCGTTTTCAGCGCATGATCCGCTGAAAG 180
QY 181 GTGAAAAACCTGCTTTATAGAGCCACCGATTTGTGAAACAGAGCATTTGACCAAGT 240
DB 181 GTGAAAAACCTGCTTTATAGAGCCACCGATTTGTGAAACAGAGCATTTGACCAAGT 240
QY 241 CTGCGGTATATACCGTATTTGAGAGCAAGTTTACTCGCGGAAAGTGAAGAAATTC 300
DB 241 CTGCGGTATATGCGGTATTTGAGAGATGCTGCTTGCAAGAGCAGCAAGAAATTC 300
QY 301 CTGCGGAAAGAGCTTGCTGCAAGAGCTTTGACCGGATACCGCTTTATGCTCGC 360
DB 301 CTGCGGAAAGATCTTGCTGCAAGAGCTTTGATAGAGATTCGCTTTATGCTCGT 360
QY 361 TTGGAACGATGTTATGCAAGTCTGACCTGCGCTCGCGGTGCGGATTTACTGCGAG 420
DB 361 TTGGAACGATGTTTTCGCAAGATTTGTCAGACCGGATTAAGTCTTAATTTAGAGAT 420
QY 421 GCGGCTTTCGCTTTGGAAGGAAACACGCGGAGGAGCGGCTATATGATTTCCGCA 480
DB 421 GCGGCTTTCGCTTTGGAAGGAAACACGCGGAGGAGCGGCTATATGATTTCCGCT 480
QY 481 AAAGGATGCGGTTTTCCTGACAGGTTTTCGCGCTGCGCGGAGGAGGCTGACCC 540
DB 481 GAGGCGATGCGGTTTTCCTGAAAGGTTTTCGCGCTGCGCGGAGGATTTAAAGC 540
QY 541 GTGCAATGAGATGTTGAGAGATTTTTCAGACAGGAGAAATCCGTTTTCAGATC 600
DB 541 GTGATGAGATGATTTTACTTATTTCTTATATGAGGAGGAGATGCTGTTATTCAGGTT 600
QY 601 AATCCGCGCTTGTGCGCCAGAGAGCTGATATGCAAGTTTCAAGCAACCAAGACGCA 660
DB 601 AATCCGCGCTTATGATACCAAGATGATATGCAAGTTTCTATGATTAAGCAAGTATG 660
QY 661 TTGGGACGCTGATGAAACAGACCGCTCTGAAACGCAACG 706
DB 661 TTGGGTAGCGATTTGAAAAAGATAGGAAACAGAAAGACAC 706
RESULT 10
ADT05534
ID ADT05534 standard; DNA; 14547 BP.
XX
AC ADT05534;
XX
DT 02-DEC-2004 (first entry)
XX
DE Haemophilus influenzae (NTHi) contig DNA sequence - SEQ ID 570.
XX
KW middle ear bacterial infection; nasopharynx bacterial infection; ds;
XX contig.
OS Haemophilus influenzae.
XX
PN WO2004078949-A2.
XX
PD 16-SEP-2004.
XX
PF 05-MAR-2004; 2004WO-US007001.
XX
PR 06-MAR-2003; 2003US-0453134P.
XX
PA (CHIL-) CHILDRENS HOSPITAL INC.

XX Bakaletz LO, Munson RS, Dyer DW;
XX WPI; 2004-662422/64.
XX
XX New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT useful for treating or preventing NTHi bacterial infections of the middle
PT ear and/or nasopharynx.
XX
XX Example 1; SEQ ID NO 570; 88bp; English.
XX
XX The invention comprises nucleotide sequences (genes) from the genome of a
CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC sequences of the invention are useful for treating or preventing NTHi
CC bacterial infections of the middle ear and/or nasopharynx. The present
CC nucleic acid represents an NTHi contig sequence of the invention.
XX
XX Sequence 14547 BP; 4491 A; 2521 C; 2901 G; 4631 T; 0 U; 3 Other;
SQ
Query Match 11.2%; Score 93; DB 13; Length 14547;
Best Local Similarity 51.9%; Pred. No. 4.7e-17;
Matches 235; Conservative 0; Mismatches 215; Indels 3; Gaps 1;
QY 180 AGTGAAGAAAGCTGCTTTATGAGCCAGCCGATTTGTGAGAGCAGCATTTGACGAAAG 239
DB 11201 AGGAGAAAGAGCTGTTTATTAAGCCATTTCTTATTAAGATTAAGTGTGAATGAAAA 11260
QY 240 TCTGCCGATATATCCAGCTATTTGAGAGCAGCTTTTACTCGCGAAGGTGAGAAAAATT 239
DB 11261 TTTGGAATATCTCAATTTTGGAGATGATGTAATTCTTGCGAGAAATGCGGAAGTGT 11320
QY 300 CTTGCCGAGAGAGCTTGCGTGCAGAAAGCTTTGACCCGAGATCCGCTTTATCGCCG 359
DB 11321 TTTGGCAAGAGATGAATGTTAAACACGTTTGATTTTAATGAATATTTTATTTATTCG 11380
QY 360 CTGGAAGACGATGTTATGACAGT---CTGACCTCGCCCTCGCGCGTGGCGGATTAAGT 416
DB 11381 TTTGAAGACTTTTAAAGCAGTAACTTGAGAAACAACTAAATTCACCTTTTAA 11440
QY 417 CGGCGCGCGCTTTCGCTGTTGAAAGCAACTGGGGGACGCGCGGCTATATCATTTTC 476
DB 11441 TTTAGAGACTTTGATATTTTAAATGACCTACCTGGGGGACGCGCTGTTATATTTTC 11500
QY 477 CCGAAAAGCGATGCGGTTTTTCTGACAGTTTCCGCGCTCGCGCGGCGAAGGCTGCA 536
DB 11501 TCAAGGTGGCGCTAAATATGTAATGAATATTTAAAGAAATATTCCTTGATGAAATTTGT 11560
QY 537 CCGCGTGCATCTGATGATGTTTCAGCGAATTTTTCGACAGGGAAGAAATGCGGTTTGCCA 596
DB 11561 TGCAGTTGATGAATCTTATTTTAAATTAATTAAGTTGATGATTAATTTATTTCTATCA 11620
QY 597 GCTCAATCCGCGCTTGTGCGCCCAAGAGCTGCA 629
DB 11621 ACTTAATCCAGCAATTTGTAATCAAGAACTCCA 11653
RESULT 11
ADT05645
ID ADT05645 standard; DNA; 106645 BP.
XX
XX ADT05645;
XX
XX 02-DEC-2004 (first entry)
XX
XX Haemophilus influenzae (NTHi) DNA sequence - SEQ ID 681.
XX
XX middle ear bacterial infection; nasopharynx bacterial infection; ds.
XX
XX Haemophilus influenzae.
XX
XX WO2004078949-A2.
XX
XX 16-SEP-2004.

XX
PF 05-MAR-2004; 2004MO-US007001.
XX
XX 06-MAR-2003; 2003US-0453134P.
XX
XX (CHIL-) CHILDRENS HOSPITAL INC.
XX
PI Bakaletz LO, Munson RS, Dyer DW;
XX
XX WPI; 2004-662422/64.
XX
XX New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT useful for treating or preventing NTHi bacterial infections of the middle
PT ear and/or nasopharynx.
XX
XX Claim 1; SEQ ID NO 681; 88bp; English.
XX
XX The invention comprises nucleotide sequences (genes) from the genome of a
CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC sequences of the invention are useful for treating or preventing NTHi
CC bacterial infections of the middle ear and/or nasopharynx. The present
CC nucleic acid represents an NTHi DNA sequence of the invention.
XX
XX Sequence 106645 BP; 32661 A; 19435 C; 21227 G; 33322 T; 0 U; 0 Other;
SQ
Query Match 11.2%; Score 93; DB 13; Length 106645;
Best Local Similarity 51.9%; Pred. No. 1.1e-16;
Matches 235; Conservative 0; Mismatches 215; Indels 3; Gaps 1;
QY 180 AGTGAAGAAAGCTGCTTTATGAGCCAGCCGATTTGTGAGAGCAGCATTTGACGAAAG 239
DB 98869 AGGAGAAAGAGCTGTTTATTAAGCCATTTCTTATTAAGATTAAGTGTGAATGAAAA 98928
QY 240 TCTGCCGATATATCCAGCTATTTGAGAGCAGCTTTTACTCGCGAAGGTGAGAAAAATT 239
DB 98929 TTTGGAATATCTCAATTTTGGAGATGATGTAATTCTTGCGAGAAATGCGGAAGTGT 98988
QY 300 CTTGCCGAGAGAGCTTGCGTGCAGAAAGCTTTGACCCGAGATCCGCTTTATCGCCG 359
DB 98989 TTTGGCAAGAGATGAATGTTAAACACGTTTGATTTTAATGAATATTTTATTTATTCG 99048
QY 360 CTGGAAGACGATGTTATGACAGT---CTGACCTCGCCCTCGCGCGTGGCGGATTAAGT 416
DB 99049 TTTGAAGACTTTTAAAGCAGTAACTTGAGAAACAACTAAATTCACCTTTTAA 99108
QY 417 CGGCGCGCGCTTTCGCTGTTGAAAGCAACTGGGGGACGCGCGGCTATATCATTTTC 476
DB 99109 TTTAGAGACTTTGATATTTTAAATGACCTACCTGGGGGACGCGCTGTTATATTTTC 99168
QY 477 CCGAAAAGCGATGCGGTTTTTCTGACAGTTTCCGCGCTCGCGCGGCGAAGGCTGCA 536
DB 99169 TCAAGGTGGCGCTAAATATGTAATGAATATTTAAAGAAATATTCCTTGATGAAATTTGT 99228
QY 537 CCGCGTGCATCTGATGATGTTTCAGCGAATTTTTCGACAGGGAAGAAATGCGGTTTGCCA 596
DB 99229 TGCAGTTGATGAATCTTATTTTAAATTAATTAAGTTGATGATTAATTTATTTCTATCA 99288
QY 597 GCTCAATCCGCGCTTGTGCGCCCAAGAGCTGCA 629
DB 99289 ACTTAATCCAGCAATTTGTAATCAAGAACTCCA 99321
RESULT 12
AAT42063_05/c
Continuation (6 of 19) of AAT42063 from base 500001 (Haemophilus influenzae complete ge
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063
WP Fragment Name Begin End
WP AAT42063_00 1 110000
WP AAT42063_01 100001 210000
WP AAT42063_02 200001 310000
WP AAT42063_03 300001 410000
WP AAT42063_04 400001 510000
WP AAT42063_05 500001 610000

WP AAT42063_06 600001 710000
 WP AAT42063_07 700001 810000
 WP AAT42063_08 800001 910000
 WP AAT42063_09 900001 1010000
 WP AAT42063_10 1000001 1110000
 WP AAT42063_11 1100001 1210000
 WP AAT42063_12 1200001 1310000
 WP AAT42063_13 1300001 1410000
 WP AAT42063_14 1400001 1510000
 WP AAT42063_15 1500001 1610000
 WP AAT42063_16 1600001 1710000
 WP AAT42063_17 1700001 1810000
 WP AAT42063_18 1800001 1930121

Query Match 10.5%; Score 86.6; DB 2; Length 110000;
 Best Local Similarity 51.0%; Pred. No. 1e-14; Mismatches 219; Indels 3; Gaps 1;
 Matches 221; Conservative 0;
 QY 180 AGTGAAGAAAGCTGCTTATGAGCCAGCCGCTATTGTGGAAGCAGGCAATTGACGAGG 239
 Db 70747 AGGAGAAAGAGGCTGTTATATGACCATTTCTATATGGAATTAAGTGTGATGAAAA 70688
 QY 240 TCTGCCGTATATACCGTATTTGAGAGCAGACCTTTTACTCGCGGAAGTGAGAAAAATT 299
 Db 70687 TTTGAAATCTCAAAATTTTGAAGATGATGTAATTTTGGCGGAATGCGGAATATT 70628
 QY 300 CCTTGCCGAAGACGCTTGCTGCAAGACGCTTGAACCCGATACCGCCTTATCGTCCG 359
 Db 70627 TTTGAACCAAAATGATGTTAAAAACAGTTTGTATTTAATGATTTTATATTTG 70568
 QY 360 CTTGGAACAGATGTTATGACGT---CCTGACCTGCCCTCGCGGTGGCGGATTACTG 416
 Db 70567 TTTAGAACTTTTTCACGCCAGTTAACTTGAAGAAACAACTTCAACTTTTAA 70508
 QY 417 CGGCGCGCTTCCGCTGTTGAAAGCAGACCTGGGAGCGCGGCTATATCATTTG 476
 Db 70507 TTTCTAGGAATTTTGAATTTTAAATCACTGCGGAGCGCAGTTATTTATTTTC 70448
 QY 477 CGGAAAGCGATGCGGCTTTTCTGTGACAGGTTTGCCTGCCGCCGGAAGGCTGCA 536
 Db 70447 TCAGAGTGGGCTTAATATGATGATTAATTAAGAAATATTTCTTCTGATGAATTTGT 70388
 QY 537 CCCCGCATCTGATGATGTTGCGGATTTTTCGACAGGAAGAAATCCCGTTTGCCA 596
 Db 70387 TGCAGTTGATGAATTTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 70328
 QY 597 GCTCAATCCGCTGTTGCGCCCAAGAGCTGCA 629
 Db 70327 ACTGAATCCAGCAATTTGTATTCAGAACTCCA 70295

RESULT 13
 ADL02801
 ID ADL02801 standard; DNA; 768 BP.
 AC ADL02801;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE DNA encoding a M. catarrhalis protein #487.
 XX
 KW ds; gene; Moraxella catarrhalis; infection.
 XX
 OS Moraxella catarrhalis.
 XX
 PN US6673910-B1.
 XX
 PD 06-JAN-2004.
 XX
 PF 04-APR-2000; 2000US-00540236.
 XX
 PR 08-APR-1999; 99US-0128416P.
 XX

PA (GENO-) GENOME THERAPEUTICS CORP.
 XX Breton GL;
 PI
 XX
 DR WPI; 2004-178127/17.
 XX F-PSDB; ADL04721.
 XX
 PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
 PT preparing a composition for diagnosing, preventing or treating infection
 PT caused by Moraxella catarrhalis.
 XX
 PS Disclosure; SEQ ID NO 487; 423bp; English.
 XX
 CC The invention relates to an isolated nucleic acid encoding a Moraxella
 CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
 CC composition for diagnosing, preventing or treating infection caused by
 CC Moraxella catarrhalis. The present sequence represents DNA encoding a M.
 CC catarrhalis protein.
 XX
 SQ Sequence 768 BP; 270 A; 142 C; 131 G; 225 T; 0 U; 0 Other;

Query Match 8.8%; Score 73.2; DB 12; Length 768;
 Best Local Similarity 52.9%; Pred. No. 1.5e-11;
 Matches 181; Conservative 0; Mismatches 158; Indels 3; Gaps 1;
 QY 1 ATGCAAAACCAAGTATGAGCTTACCTCCGCGAGAACGACGAGCGCACATTGCCGAT 60
 Db 7 ATACAAAATTTTGTATAGTGTAAATCTGCCAAGAAAGAAAGACATTTATGTGT 66
 QY 61 ACCTTGCGGAGGACGCGATCCCGTTGAGCTTTTTCAGCACTGATGCCGTGAAAGG 120
 Db 67 GAATTTGGCAACAGAGGATTTGCTTTGATGAGTAAACCCACCTGATTTT 126
 QY 121 CTGGAACAGCAATGCGGAATCGTCCCGCTTGTGGCGCACCCCTATTGAGCGGA 180
 Db 127 AGCAAGTACGCCCAAAAGCTTTCATCCCATTCATCAACA---ACCAAGACTCACCGAT 183
 QY 181 GTGAAAAAGCTGCTTATATGAGCCAGCGCTATTGTGGAAGCAGGCAATTGACGAGGT 240
 Db 184 GCGCAAAAAGCTGCTTATTAAGCCATGTACATATGCAACAGATGATGATGAAAAAC 243
 QY 241 CTGCGCTATATCACCGTATTTGAGAGCAGCTTTTACTCGCGGAGGAGGAAAAATTC 300
 Db 244 TTGATTTATATGCAATTTTGAAGCAGATGCTTATTTGGGAATGACTCTCAACAATTC 303
 QY 301 CTGCGGAAGACGCTTGCTGCAAGAACGCTTTGACCCGAT 342
 Db 304 TTACAAAGAACTGACAAATTTGGCTACAACAAAATGCGGTTGAT 345

RESULT 14
 ADV09334/C
 ID ADV09334 standard; DNA; 4583 BP.
 AC ADV09334;
 XX
 DT 24-FEB-2005 (first entry)
 XX
 DE M. catarrhalis strain 7169 serotype B LOS gene cluster.
 XX
 KW ds; diagnosis; bacterial infection; otitis media; LOS gene cluster.
 XX
 OS Moraxella catarrhalis.
 XX
 PN W02004104170-A2.
 XX
 PD 02-DEC-2004.
 XX
 PF 13-MAY-2004; 2004WO-US015141.
 XX
 PR 13-MAY-2003; 2003US-0470022P.
 XX
 PA (UNIV) UNIV NEW YORK STATE RES FOUND.

PA (LUCK/) LUCK N.
PA (FURA/) FURANO K.
PA (HOWL/) HOWLETT A.
XX
PI Campagnari AA;
XX
DR WPI; 2005-013280/01.
XX
XX
PT Detecting the presence of Moraxella catarrhalis in acute otitis media by
PT detecting the presence of proteins, nucleic acid sequences and
PT lipooligosaccharides (LOS) specific to M. catarrhalis.
PS
XX
XX Example 5; SEQ ID NO 16; 52pp; English.
XX
CC This invention describes a novel method for detecting the presence of
CC Moraxella catarrhalis in a biological sample and comprises contacting the
CC biological sample with at least two antibodies chosen from 7c9, 4g5, 3f7
CC or 3f5, where the epitope in M. catarrhalis for each antibody is either
CC internal or surface exposed. The invention also describes a method of
CC identifying and determining M. catarrhalis lipooligosaccharide (LOS)
CC serotypes in a biological sample using monoclonal antibodies which
CC recognize M. catarrhalis serotypes A, B and C or by amplifying M.
CC catarrhalis genomic DNA with PCR primers. The detected gene is a gene
CC which encodes a protein or epitope that is unique to Moraxella
CC catarrhalis LOS A, LOS B, LOS C and is either exposed on the surface of
CC M. catarrhalis (see ADV09320) or not expressed on the surface of
CC Moraxella catarrhalis (see ADV09324). The gene is detected by PCR
CC amplification. The methods and compositions of the present invention are
CC useful for providing tools for specifically detecting M. catarrhalis
CC particularly in bacteria-induced acute otitis media. This sequence
CC represents the M. catarrhalis strain 7169 type B LOS gene cluster which
CC can be detected by PCR amplification using PCR primers 406 and 408
CC (ADV09325 & ADV09326).
XX
XX
SQ Sequence 4583 BP; 1346 A; 853 C; 931 G; 1453 T; 0 U; 0 Other;
Query Match 8.8%; Score 73.2; DB 14; Length 4583;
Best Local Similarity 52.9%; Pred. No. 3.2e-11;
Matches 181; Conservative 0; Mismatches 158; Indels 3; Gaps 1;
QY 1 ATGCAGAACCAAGCTTATGAGCTTACGCTTCCGCGGAGAAAGCGAGCGGACATTTGCCGAT 60
DB 2031 ATACAAAATTTTGTATCATGATGTAATAAATCGCCACAAAAGAGAAACATATTATGTGT 1972
QY 61 ACCCTGCGAGGAGCGGATCCCGTTTCACTTTTGAAGCACTAGTCCGCTGAAAG 120
DB 1971 GAATTTGGCAAAACAAAGCATGTGCTTTGAGTTTGTGATGACAGTAAACCCCATGTATTT 1912
QY 121 CTGGAACAGGCAATGCGGAACCTGTCGCCGCTGTGCGCGCAACCCCTATTTTGACGGA 180
DB 1911 AGCAAGTAGCCCAAAAGCTTCAATCCCAATCA---CCAAACACCAAAAGATCACCAGAT 1855
QY 181 GTGGAAAAAGCTGCTTTATGAGCCAGCGGATTTGTGAGAGGAGCATTTGACGAAGGT 240
DB 1854 GCGGAAAAAGCTGCTTTTAAAGCCATGTAGCACTATGCGAACAAGATGATGTAAGAAAC 1795
QY 241 CTGCGCTATATCAACCGTATTTGAGAGCAAGCTTTTACTCGGCGAAGAGTGAAGAAATTC 300
DB 1794 TTGATTAATATGCAATTTTGTGAGAGCAATGCTATTTTGGGGAATACCTCTCAACAAATTC 1735
QY 301 CTGCGGAGAGCGCTTGCGTGAAGAAGCTTTGACCCGAGT 342
DB 1734 TTACAAAGAACTGACAAATTTGGCTACAACAAAATGCCGTTGAT 1693

RESULT 15
ADV09335/c
ID ADV09335 strand; DNA; 5584 BP.
XX
AC ADV09335;
XX
XX 24-FEB-2005 (first entry)
XX

DE M. catarrhalis strain RS-10 serotype C LOS gene cluster.
XX
XX de; diagnosis; bacterial infection; otitis media; LOS gene cluster.
XX
XX Moraxella catarrhalis.
XX
OS
XX
PN MO2004104170-A2.
XX
XX
PD 02-DGC-2004.
XX
XX
PF 13-MAY-2004; 2004WO-US015141.
XX
XX
PR 13-MAY-2003; 2003US-0470022P.
XX
XX
PA (UNNY) UNIV NEW YORK STATE RES FOUND.
PA (LUCK/) LUCK N.
PA (FURA/) FURANO K.
PA (HOWL/) HOWLETT A.
XX
PI Campagnari AA;
XX
DR WPI; 2005-013280/01.
XX
XX
PT Detecting the presence of Moraxella catarrhalis in acute otitis media by
PT detecting the presence of proteins, nucleic acid sequences and
PT lipooligosaccharides (LOS) specific to M. catarrhalis.
PS
XX
XX Example 5; SEQ ID NO 17; 52pp; English.
XX
CC This invention describes a novel method for detecting the presence of
CC Moraxella catarrhalis in a biological sample and comprises contacting the
CC biological sample with at least two antibodies chosen from 7c9, 4g5, 3f7
CC or 3f5, where the epitope in M. catarrhalis for each antibody is either
CC internal or surface exposed. The invention also describes a method of
CC identifying and determining M. catarrhalis lipooligosaccharide (LOS)
CC serotypes in a biological sample using monoclonal antibodies which
CC recognize M. catarrhalis serotypes A, B and C or by amplifying M.
CC catarrhalis genomic DNA with PCR primers. The detected gene is a gene
CC which encodes a protein or epitope that is unique to Moraxella
CC catarrhalis LOS A, LOS B, LOS C and is either exposed on the surface of
CC M. catarrhalis (see ADV09320) or not expressed on the surface of
CC Moraxella catarrhalis (see ADV09324). The gene is detected by PCR
CC amplification. The methods and compositions of the present invention are
CC useful for providing tools for specifically detecting M. catarrhalis
CC particularly in bacteria-induced acute otitis media. This sequence
CC represents the M. catarrhalis strain RS-10 type C LOS gene cluster which
CC can be detected by PCR amplification using PCR primers 406, 408, 750 and
CC 751 (ADV09325, ADV09326, ADV09329 and ADV09330).
XX
XX
SQ Sequence 5584 BP; 1664 A; 1048 C; 1135 G; 1737 T; 0 U; 0 Other;
Query Match 8.8%; Score 73.2; DB 14; Length 5584;
Best Local Similarity 52.9%; Pred. No. 3.5e-11;
Matches 181; Conservative 0; Mismatches 158; Indels 3; Gaps 1;
QY 1 ATGCAGAACCAAGCTTATGAGCTTACGCTTCCGCGGAGAAAGCGAGCGGACATTTGCCGAT 60
DB 3028 ATACAAAATTTTGTATCATGATGTAATAAATCGCCACAAAAGAGAAACATATTATGTGT 2969
QY 61 ACCCTGCGAGGAGCGGATCCCGTTTCACTTTTGAAGCACTAGTCCGCTGAAAG 120
DB 3028 ATACAAAATTTTGTATCATGATGTAATAAATCGCCACAAAAGAGAAACATATTATGTGT 2969
QY 2968 GAATTTGGCAAAACAAAGCATGTGCTTTGAGTTTGTGATGACAGTAAACCCCATGTATTT 2909
QY 121 CTGGAACAGGCAATGCGGAACCTGTCGCCGCTGTGCGCGCAACCCCTATTTTGACGGA 180
DB 2908 AGCAAGTAGCCCAAAAGCTTCAATCCCAATCAACA---ACCAAAAGATCACCAGAT 2852
QY 181 GTGGAAAAAGCTGCTTTATGAGCCAGCGGATTTGTGAGAGGAGCATTTGACGAAGGT 240
DB 2851 GCGGAAAAAGCTGCTTTTAAAGCCATGTAGCACTATGCGAACAAGATGATGTAAGAAAC 2792
QY 241 CTGCGCTATATCAACCGTATTTGAGAGCAAGCTTTTACTCGGCGAAGAGTGAAGAAATTC 300

Db 2791 TTGGATTATATGGCAATTTTGAAGACGATGTCTATTGGGAATGACTCTCAACAATTC 2732
Qy 301 CTTGCCGAAGACGCTTGGCTGCAAGAACGCTTTGAACCCGAT 342
Db 2731 TTACAAGAACTGACAATTTGGCTACAAACAATAAGCCGTTGAT 2690

Search completed: April 7, 2006, 16:17:30
Job time : 938.991 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:00:12 ; Search time 6502.61 Seconds
(without alignments)
5957.564 Million cell updates/sec

Title: US-09-211-691-1

Perfect score: 828
Sequence: 1 atgcacaaacacgtatcacg.....tcattgtgccttccaataa 828

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40.2	4.9	522	6	CD231899 SS1_30_A0
2	40.2	4.9	598	6	CD231864 SS1_30_A0
3	40.2	4.9	620	6	CB926280 ABA1_7_H0
4	40.2	4.9	729	6	CP430189 P11_36_H0
5	39.2	4.8	1217	8	DR135967 43289438
6	39.2	4.7	530	1	AA461978 V971F07.r
7	39.2	4.7	1357	9	CC306269 CH261-2K1
8	39.2	4.7	737	9	CC821108 FPCPA98 U
9	38.6	4.7	582	3	BP210238 BP210238
10	38.6	4.7	469	1	AM677069 DGL_4_C12
11	38.6	4.7	775	9	BZ577887 msh2_5621
12	38.4	4.6	920	4	CR651152 tetradon
13	38.2	4.6	569	7	CN206175 T0r607 G
14	38.2	4.6	553	6	CF687694 CCACV03TR
15	38.2	4.6	591	6	CF712358 CCAB080TR
16	38.2	4.6	597	6	CF693664 CCAM50TR
17	38.2	4.6	619	8	DR826908 ZM_BFB006
18	38.2	4.6	666	6	CF678610 CCAD48TR
19	38.2	4.6	693	6	CF709985 CCABU50TR
20	38.2	4.6	715	6	CF695835 CCABU76TR
21	38.2	4.6	735	6	CF679141 CCABF28TR
22	38.2	4.6	743	6	CF713838 CCAL52TR

23	38	4.6	743	6	CF721366 CCAX90TR
24	38	4.6	754	6	CF714073 CCABU78TR
25	38	4.6	749	8	DR954928 ZM_BFB004
26	38	4.6	755	6	CF696727 CCABP79TR
27	38	4.6	757	6	CF675743 CCAD07TR
28	38	4.6	760	9	BZ538612 OGAI87TC
29	38	4.6	762	6	CF678653 CCABH89TR
30	38	4.6	763	7	CO526725 CCABV1.1.17
31	38	4.6	779	6	CF701635 CCAS522TR
32	38	4.6	781	6	CF676109 CCABH75TR
33	38	4.6	782	6	CF688826 CCAD047TR
34	38	4.6	789	6	CF690987 CCADK83TR
35	38	4.6	793	6	CF676377 CCABH89TR
36	38	4.6	812	6	CF719729 CCACK73TR
37	38	4.6	816	6	CF682178 CCAD013TR
38	38	4.6	827	6	CF676273 CCAD013TR
39	38	4.6	829	6	CF712852 CCADH51TR
40	38	4.6	829	6	CF718427 CCAGH17TR
41	38	4.6	859	6	CF698378 CCAP057TR
42	38	4.6	862	6	CF696539 CCABH93TR
43	38	4.6	863	6	CF696079 CCAGB95TR
44	38	4.6	870	6	CF696612 CCAD102TR
45	38	4.6	872	6	CF683469 CCACW55TR

ALIGNMENTS

RESULT 1
CD231899/c
LOCUS
DEFINITION
SS1_30_A04_g1_A012 Salt-stressed seedlings Sorghum bicolor cDNA
Clone SS1_30_A04_A012 5', mRNA sequence.
CD231899
CD231899.1 GI:30975364
VERSION
KEYWORDS
SOURCE
ORGANISM
Sorghum bicolor (sorghum)
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 522)
Cordonier-Pratt, M.-M., Wentzel, V., Suzuki, Y., Sugano, S.,
Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Summer, E.J.,
Bastman, A. and Pratt, L.H.
An EST database from Sorghum: salt-stressed seedlings
Unpublished (2003)
Other ESTs: SS1_30_A04_b1_A012
Contact: Cordonier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug5 (CTTCTGCTCTAAAGCTGCG).
Location/Qualifiers

FEATURES

source
1..522
/organism="Sorghum bicolor"
/mol_type="mRNA"
/catalytic="133620C"
/db_xref="taxon:4558"
/clone="SS1_30_A04_A012"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Salt-stressed seedlings"

ORIGIN

/note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from 9-day-old seedlings grown in hydroponic culture. Seedlings were transferred to a 150 mM NaCl solution and harvested at 3, 6, 12 and 24 hr following transfer. Roots and leaves were pooled from all time points and RNA isolated. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5'-prime DraIII site is CACTGTGTG, 3'-prime DraIII site is CACCATGTG)."

Query Match 4.9%; Score 40.2; DB 6; Length 522;
Best Local Similarity 49.3%; Pred. No. 2.2;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 215 TGTGAGAGGAGGATTTGACGAGAGTCTCCGTATATACCGTATTTGAGCAGACGTTT 274
DB 415 TCTGGAAGCTGTTTGGGAGACGAGAACTGATGTTCCAAATGTTCTGAGGAGCACTAACT 356
QY 275 TACTCGGCGAAGGTGAGAAAATTCTTGCAGAAAGCGTTGCGTGCAGAAAGCGTTTG 334
DB 355 TGTCTCGAAGCGGCTCTTAACAAGCTTTGCCGATGTCCTTGAGCCGAGAACTCCTCTT 296
QY 335 ACCCGAATACCGCCTTTATGCTTCGCTTGAAGAAAGATTTATGCAAGTCTGACCTGCG 394
DB 295 AGCAGCAGAGTGTCTGCTTGTGAGCGGAGCGAGAGTCTGCGCGTGCAGAGCCCGCA 236
QY 395 CCTCCGCGGTGCGGATTAATGCGGCGCGCT 427
DB 235 GCGGCGCAGGAGGAGTGTGCGGTTGAGAGCAGCT 203

RESULT 2
LOCUS CD231864 598 bp mRNA linear EST 21-MAY-2003
DEFINITION SSI_30_A08_g1_A012 salt-stressed seedlings Sorghum bicolor CDNA
ACCESSION CD231864
VERSION CD231864.1 GI:30975329
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor

REFERENCE 1 (bases 1 to 598)
AUTHORS Cordonnier-Pratt, M.-M., Wentzel, V., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Summer, E.J., Eastman, A. and Pratt, L.H.
COMMENT An EST database from Sorghum: salt-stressed seedlings

Unpublished (2003)
Other ESTs: SSI_30_A08_b1_A012
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210

Email: mmpat@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

Seg primer: Sug5 (CTTCTGCTCTAAAGCTGCG).

FEATURES
source 1..598
Location/Qualifiers
/organism="Sorghum bicolor"
/mol_type="mRNA"

ORIGIN

/cultivar="HS3620C"
/db_xref="taxon:4558"
/clone="SSI_30_A08_A012"
/lab_host="DH10B-rT phage-resistant E. coli"
/clone_id="Salt-stressed seedlings"
/note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from 9-day-old seedlings grown in hydroponic culture. Seedlings were transferred to a 150 mM NaCl solution and harvested at 3, 6, 12 and 24 hr following transfer. Roots and leaves were pooled from all time points and RNA isolated. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5'-prime DraIII site is CACTGTGTG, 3'-prime DraIII site is CACCATGTG)."

Query Match 4.9%; Score 40.2; DB 6; Length 598;
Best Local Similarity 49.3%; Pred. No. 2.3;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 215 TGTGAGAGGAGGATTTGACGAGAGTCTCCGTATATACCGTATTTGAGCAGACGTTT 274
DB 426 TCTGGAAGCTGTTTGGGAGACGAGAACTGATGTTCCAAATGTTCTGAGGAGCACTAACT 367
QY 275 TACTCGGCGAAGGTGAGAAAATTCTTGCAGAAAGCGTTGCGTGCAGAAAGCGTTTG 334
DB 366 TGTCTCGAAGCGGCTCTTAACAAGCTTTGCCGATGTCCTTGAGCCGAGAACTCCTCTT 307
QY 335 ACCCGAATACCGCCTTTATGCTTCGCTTGAAGAAAGATTTATGCAAGTCTGACCTGCG 394
DB 306 AGCAGCAGAGTGTCTGCTTGTGAGCGGAGCGAGAGTCTGCGCGTGCAGAGCCCGCA 247
QY 395 CCTCCGCGGTGCGGATTAATGCGGCGCGCT 427
DB 246 GCGGCGCAGGAGGAGTGTGCGGTTGAGAGCAGCT 214

RESULT 3
LOCUS CB926280 620 bp mRNA linear EST 28-APR-2003
DEFINITION ABA1_7_H09_g1_A012 Abscisic acid-treated seedlings Sorghum bicolor CDNA clone ABA1_7_H09_A012 5', mRNA sequence.
ACCESSION CB926280
VERSION CB926280.1 GI:30162551
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor

REFERENCE 1 (bases 1 to 620)
AUTHORS Cordonnier-Pratt, M.-M., Wentzel, V., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Buchanan, C.D., Eastman, A. and Pratt, L.H.
COMMENT An EST database from Sorghum: ABA1-treated seedlings

Unpublished (2003)
Other ESTs: ABA1_7_H09_b1_A012
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210

Email: mmpat@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

FEATURES
source 1..598
Location/Qualifiers
/organism="Sorghum bicolor"
/mol_type="mRNA"

FEATURES
Seq primer: Sug5 (CTTCTGCTCTAAGAGTCGC).
Location/Qualifiers
1. .620
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone="ABAI 7 H09 A012"
/lab_host="DH10B-T1 phage-resistant E. coli"
/note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from seedlings grown in hydroponic culture. After 12 days, medium was supplemented with 1 mM abscisic acid (ABA), while leaves were misted with a solution of 1 mM ABA. Roots and leaves were harvested after 3, 6, 12, and 24 hr and material from all time points was combined prior to RNA isolation. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

ORIGIN

Query Match 4.9%; Score 40.2; DB 6; Length 620;
Best Local Similarity 49.3%; Pred. No. 2.3;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 215 TGTGGAAGCAGCATTTGAGAGAGGTCGCCATATACCGCTATTGAGAGACGCTTT 274
DB 446 TCTGGAAGCTGTTTGGGAGACGAGAACTGATGTTCCATGTTCTGAGACCATTAAT 387
QY 275 TACTCGCGAAGGTGAGAAAAATTCCTGCGGAGACGCTGCTGCAAGAGCGCTTG 334
DB 386 TGTCTCGAAGCGGTCTCTAACAAGCTTTCCTGATGCTTGAGCCAGAAATCGCTTT 327
QY 335 ACCCGATACCGCCTTATGTCGCTTGAGAAAGATGTTTATGACGTCGACCTCGC 394
DB 326 AGCAGCAGAGTGTCTGCTGTGAGCGGCTGCGAAGTCGTGCGGTGAGAGCCCGCA 267
QY 395 CTTCCGCGGTGCGGATTAATCGCGGCGCGCT 427
DB 266 GCGCGCGCAGGAGGTGCGGCTGAGGAGCT 234

RESULT 4
LOCUS CP430189/c 729 bp mRNA linear EST 03-SEP-2003
DEFINITION PH1_26_H07_g1_A002 Phosphorous-deficient seedlings Sorghum bicolor
ACCESSION CP430189
VERSION CP430189.1 GI:34442890
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Bukariyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 729)
Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R. R., Liang, C., Sun, F., Sullivan, R., Harris, K., Eastman, A. and Pratt, L. H.
An EST database from Sorghum: phosphorous-deficient seedlings
Unpublished (2003)
Other ESTs: PH1_26_H07_b1_A002
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University;

sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
Seq primer: Sug5 (CTTCTGCTCTAAGAGTCGC).
Location/Qualifiers
1. .729
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone="PH1_26_H07 A002"
/lab_host="DH10B-T1 phage-resistant E. coli"
/note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from Brix623 seedlings grown hydroponically in the absence of added phosphorous. At 14 days of age, roots and shoots were harvested and stored at -80 C until RNA was isolated. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

ORIGIN

Query Match 4.9%; Score 40.2; DB 6; Length 729;
Best Local Similarity 49.3%; Pred. No. 2.4;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 215 TGTGGAAGCAGCATTTGAGAGAGGTCGCCATATACCGCTATTGAGAGACGCTTT 274
DB 431 TCTGGAAGCTGTTTGGGAGACGAGAACTGATGTTCCATGTTCTGAGACCATTAAT 372
QY 275 TACTCGCGAAGGTGAGAAAAATTCCTGCGGAGACGCTGCTGCAAGAGCGCTTG 334
DB 371 TGTCTCGAAGCGGTCTCTAACAAGCTTTCCTGATGCTTGAGCCAGAAATCGCTTT 312
QY 335 ACCCGATACCGCCTTATGTCGCTTGAGAAAGATGTTTATGACGTCGACCTCGC 394
DB 311 AGCAGCAGAGTGTCTGCTGTGAGCGGCTGCGAAGTCGTGCGGTGAGAGCCCGCA 252
QY 395 CTTCCGCGGTGCGGATTAATCGCGGCGCGCT 427
DB 251 GCGCGCGCAGGAGGTGCGGCTGAGGAGCT 219

RESULT 5
LOCUS DR135967/c 1217 bp mRNA linear EST 16-JUN-2005
DEFINITION 49289438 Drosophila pseudoobscura embryonic cDNA library Drosophila
ACCESSION DR135967
VERSION DR135967.1 GI:67881057
KEYWORDS EST.
SOURCE Drosophila pseudoobscura
ORGANISM Drosophila pseudoobscura
Bukariyota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1217)
Richard, S., Liu, Y., Bettencourt, B. R., Hradecky, P., Ietovsky, S., Nielsen, R., Thornton, K., Hudis, M. J., Chen, R., Weisel, R. P., Coutome, O., Hua, S., Smith, M. A., Zhang, P., Liu, J., Buessemaker, H. J., van Batenburg, M. F., Howells, S. L., Scherer, S. E., Sodergren, E., Matthews, B. B., Crosby, M. A., Schroeder, A. J., Ortiz-Barrientos, D., Rives, C. M., Metzger, M. L., Muzny, D. M., Scott, G., Steffen, D., Wheeler, D. A., Worley, K. C., Havlak, P., Durbin, K. J., Egan, A., Gill, R., Hume, J., Morgan, M. B., Miner, G., Hamblon, C., Huang, Y., Waldron, L., Verdusco, D., Clerc-Blankenburg, K. P., Dubchak, I., Noor, M. A., Anderson, W., White, K. P., Clark, A. G., Schaeffer, S. W., Gelbart, W., Weinstein, G. M. and Gibbs, R. A.
Comparative genome sequencing of Drosophila pseudoobscura;

TITLE

FEATURES High quality sequence stop: 615.

Location/Qualifiers

1..1357
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-2K1"
/sex="female"
/cell_line="HCD001, inbred 256"
/clone_1ib="CH261"
/note="Vector: pPARBAC1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: <http://www.chori.org/bacpac>"

ORIGIN

Query Match 4.7%; Score 39.2; DB 9; Length 1357;
Best Local Similarity 51.7%; Pred. No. 5.7;
Matches 89; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

625 CTGCATTATGCGCAAGTTTCAGACCAAAACAGCGCATTTGGGCGCCTGATCGAACACGAC 684
|||
1059 CTCTCTTACCACTCTCCAGACCAAGAAACCAACCAATTCTTGGCCCAACCAACAA 1118
|||
685 CGCCTCTGAAACCGCAACGACAAAGCGCGATTTCCCGCGCAACATTCATTAACACCGC 744
|||
1119 CGCGCTACCAACCAAAACAGACAGCGCCCTCCCTCCCAAAAATACTACAGAACAC 1178
|||
745 CTGATCCGCGCCTTGAACCAATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 796
|||
1179 AAACACCCACCATATACCAACCAAGCGCGCAAAATACCAACCA 1230
|||

RESULT 8
CC821108/c 737 bp DNA linear GSS 22-OCT-2003
DEFINITION FCBP48 Uncultured human fecal virus uncultured human fecal virus
ACCESSION CC821108
VERSION CC821108.1 GI:37806908
KEYWORDS GSS.
SOURCE uncultured human fecal virus
ORGANISM uncultured human fecal virus
REFERENCE 1 (bases 1 to 737)
Breitbart, M., Hewson, I., Felts, B., Mahaffy, J. M., Nulton, J.,
Salomon, P., and Rohwer, F.
Metagenomic analyses of an uncultured viral community from human
feces

JOURNAL J. Bacteriol. 185 (20), 6220-6223 (2003)

COMMENT
Contact: Rohwer F
Biology Dept.
San Diego State University
5500 Campanile Dr., San Diego, CA 92102, USA
Tel: 6195941336
Fax: 619595676
Email: foreste@umstroke.sdsu.edu
Classes: Shocgun.

FEATURES
source
1..737
Location/Qualifiers

/organism="uncultured human fecal virus"
/mol_type="genomic DNA"
/db_xref="taxon:239364"
/clone_1ib="Uncultured human fecal virus"

ORIGIN

Query Match 4.7%; Score 39; DB 9; Length 737;
Best Local Similarity 49.3%; Pred. No. 5.5; Indels 0; Gaps 0;
Matches 102; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

262 GAGGACGAGCTTTACTCGCGGAGGTGAGAAATTCCTTCCGAGACGCTTGCTG 321
|||

Db 605 GAGTACGCGGAGAGTGCGGAGACGAGATACAGAGACATTTAGAGGCCAGGCGCG 546
|||
Qy 322 CAAGAACGCTTGAACCCGATACCGCTTTATCGTCCGTTGGAACGATGTTATGAC 381
|||
Db 545 GAAATTAACCTGGCGCGGATGATGCTTTTCAACTTGTGAGGAGGAGATGAG 486
|||
Qy 382 GTTCCTGACCTCGCCCTCCGCGTGGCGGATTAATGCGGCGCGCTTTCGCTTGA 441
|||
Db 485 GACCACTACTCCCCCGCGCGGCTGCGCTTACATGAGAGGCGCGCTGAGTTGA 426
|||
Qy 442 AGGAACTGCGGAGCGCGGCTAT 468
|||
Db 425 ACCCACTCTCGGAAACCACTTAT 399
|||

RESULT 9
BP210238 582 bp mRNA linear EST 15-SHP-2004
LOCUS BP210238 Sugano cDNA library, cerebellum Homo sapiens cDNA clone
DEFINITION CB105056, mRNA sequence.
ACCESSION BP210238
VERSION BP210238.1 GI:52083129
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 582)
Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
15342556

JOURNAL J. Neurosci. 24 (36), 8000-8008 (2004)
COMMENT
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokane-dai, Minato-ku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..582
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CB105056"
/issue_type="cerebellum"
/clone_1ib="Sugano cDNA library, cerebellum"

ORIGIN

Query Match 4.7%; Score 38.8; DB 3; Length 582;
Best Local Similarity 45.4%; Pred. No. 5.9;
Matches 139; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

411 TTACGCGCGCGCGCTTTCCGCTTTGGAAGCAACATCGGGGAGCGCGGCTATAT 470
|||
Db 5 TTGATGAGGAGAGAGCTCACTGATGATGAGACCAACCTCGGGGGGTCAAGGCTGT 64
|||
Qy 471 CATTTCGCGAAGAGAGCGGTTTCTCTGACAGTTTCCGCGCGCTCCCGCGGAAG 530
|||
Db 65 CTTTCTTCCCTCCGCTCGCTCGCGCGCTCCACACAGTTGCAACCTGCAAGAGCGCGGAGA 124
|||
Qy 531 GCTGACCCCGCTGATCTGATGATTTTCAAGGATTTTTCAGACGAGAGAGATCCGCT 590
|||
Db 125 ACACAACTCTCCGAGAGCCAGGTCAGAGCCAAACCGTCACTGACAGTTGGCTGA 184
|||
Qy 591 TTGCAAGCAATCCCGCTTGTGCGCCAGAGCTGATTAATGCAAGTTTCAAGACA 650
|||
Db 185 TGATGCGCTGGAAGACCGCTGATGCTGCACTGTTAAGAGATTAAGATGCTGG 244
|||
Qy 651 AAACAGCCATTTGGGACCTGATGAAACAGACCGCTCTGAAACCCGAAACAGCAAG 710
|||
Db 245 TATCAAGACTTTGGTTATGTTAGTGAACAGAGGAAACAATGGAACGATTTGAGGAGG 304
|||

```

QY      711 GCGCGA 716
DB      305 GATGGA 310

RESULT 10
LOCUS   AM677069/c
DEFINITION DGI_4_C12_b1_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA
ACCESSION AM677069
VERSION   AM677069
KEYWORDS  EST.
SOURCE    Sorghum bicolor (sorghum)
ORGANISM  Sorghum bicolor (sorghum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 469)
Cordonnier-Pratt,M.-W., Gingle,A., Marsala,C., Sudman,M. and
Pratt,L.H.
An EST database from Sorghum: dark-grown seedlings
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 412
POLYA=No.
FEATURES
    source
        1..469
            /location/Qualifiers
            /organism="Sorghum bicolor"
            /mol_type="mRNA"
            /db_xref="taxon:4558"
            /clone_11b="Dark Grown 1 (DGI)"
            /note="Organ: 5-day-old dark-grown seedlings; Vector:
            lambda Zap; Site 1: XhoI; Site 2: EcoRI; The library was
            made from poly-A RNA in the cloning vector lambda Zap II.
            Clones to be sequenced were prepared by mass excision."
ORIGIN
Query Match      4.7%; Score 38.6; DB 1; Length 469;
Best Local Similarity 49.7%; Pred. No. 6.4;
Matches 98; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY      215 TGTGAGAGGAGGATTTGAGCAAGGTCGCGTATATACCGTATTGAGCAGCGTTT 274
DB      205 TGTGAGAGGAGGATTTGAGGAGCAGAACTGATGTTCCATGTTTTCGAGGAGCCATACT 146
QY      275 TACTCGGCGAGGTGAGAAAATTCCTTGCAGAGACGCTTGCGTGAAGACCGCTTGG 334
DB      145 TGCTCCGAGACCGGTCCTTAACAAGCTTGGCGATGTCGTTAGGCCCAAGATCCCTCTT 86
QY      335 ACCCGGATTCGCGCTTTATCGTCGCGCTTGGAAGAGATGTTATGACGCGCTGACCTGCG 394
DB      85 AGCAGCAGAGTTCCTGCTTTCGAGCGGGGTGGCGGAAGTCCTCGCGTCGAGGCCCGGA 26
QY      395 CCTCGGCGGTGGCGGAT 411
DB      25 GCGGCCGCGAGGAGGT 9

RESULT 11
LOCUS   BZ577887/c
FEATURES
    source
        775 bp
        DNA
        linear
        GSS 17-DEC-2002

```

```

DEFINITION msh2_5621.x1 msh Pseudomonas aeruginosa genomic clone msh2_5621,
genomic survey sequence.
ACCESSION BZ577887
VERSION   BZ577887.1
KEYWORDS  GSS.
SOURCE    Pseudomonas aeruginosa
ORGANISM  Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 775)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 357145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
    source
        1..775
            /location/Qualifiers
            /organism="Pseudomonas aeruginosa"
            /mol_type="genomic DNA"
            /strain="MSH"
            /db_xref="taxon:287"
            /clone_11b="msh2_5621"
            /note="Environmental isolate. Whole genomic shotgun
            library."
ORIGIN
Query Match      4.7%; Score 38.6; DB 9; Length 775;
Best Local Similarity 49.3%; Pred. No. 7.3;
Matches 101; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY      344 CCGCCTTATGTCGCGCTTGGAAGAGATGTTATGCAAGTCGACCTCGCCCTCCGCGC 403
DB      546 CCGCGTATGCGCTTGAGGCTTGAGCTTGCGCGCGCGCTGCGCAGAGTCAACGG 487
QY      404 TGGCGATTTACTGCGGCGCGCCCTTTCGCTGTGGAAGCGAACACTGCGGCGCGCG 463
DB      486 TGCTGAGTACTCTGAGGAGATGTCGCCGCGGCTGAGATGGGAAACCGAAGACCCCTGC 427
QY      464 GCTATATATTTCCCGAAAGCGATGCGGTTTTTCTTGACAGAGTTTGCCTGCGCGC 523
DB      426 AGCGCGCTTACCCCGCAGGCGCATGCGGTTTCGCTCGCACCCGCGTGCCTCGCGC 367
QY      524 CCGAAGGCGTCACCCGCTGATCT 548
DB      366 GCATGCGGAGCAGGCGCTCGAAT 342

RESULT 12
LOCUS   CNSOP4HI
DEFINITION Tetradon nigriviridis full-length cDNA.
ACCESSION CR655152
VERSION   CR655152.2
KEYWORDS  HTC; cDNA; full-length; full-length cDNA; Tetradon nigriviridis.
SOURCE    Tetradon nigriviridis
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1
Jaillon,O., Aury,J.M., Brunet,F., Petit,J.L., Stange-Thomann,N.,
Mauceil,E., Bouneau,L., Fischer,C., Ozouf-Costaz,C., Bernot,A.,
Nicaud,S., Jaffe,D., Fisher,S., Lutfalla,G., Dossat,C.,

```

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 JOURNAL
 TITLE
 COMMENT
 FEATURES
 SOURCE
 ORIGIN
 Query Match
 Best Local Similarity 51.1%; Score 38.4; Length 920;
 Matches 90; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
 Oy 631 TATGCCAAGTTTACGACCAAAACGCGCATTTGGCAGCTGATGAAACGACCGCTTC 690
 Db 236 TCTGCTGGAGAAACGAGAAAGGCGGAAAGGGCGACGGTACAGCTGATATCATC 295
 Oy 691 CTGAACGCGCAACGCAAGAGGCGGATTCCTCCCGCCACACATTCACACACCGCTGATC 750
 Db 296 CTCACGCAACCTGCAACCTCTCCGCTTTCTCAAAAAGTCTTATGCTTACGTCAA 355
 Oy 751 CGCGCTTGCACCAAAATCAGCAGGAAAGGAAACCGCGGCAAAAGCGCGCAACA 806
 Db 356 GACTACATGAAAGCATCATAGAGCCACTCGAGAGAAACCAACCCAGCGCGCATCA 411
 RESULT 13
 CN206175 569 bp mRNA linear EST 30-APR-2004
 LOCUS CN206175
 DEFINITION Tor6607 Gametophyte rehydration Library Tortulla ruralis cDNA, mRNA
 sequence.
 ACCESSION CN206175
 VERSION CN206175.1 GI:46902906
 KEYWORDS EST.
 SOURCE Tortulla ruralis
 ORGANISM Tortulla ruralis
 Buariyoda; Vitridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Dictyotales; Pottiales; Pottiaceae; Tortulla.
 1 (bases 1 to 569)
 Oliver,M.J., Dowd,S.E., Zaragosa,J., Mauge,S.A. and Payton,P.R.
 The rehydration transcriptome of the desiccation-tolerant bryophyte
 Tortulla ruralis: transcript classification and analysis
 BMC Genomics 5 (1), 89 (2004)
 15546486
 Contact: Oliver Melvin J
 Plant Stress Lab
 USDA-ARS
 3810 4th St, Lubbock, TX 79415, USA
 Tel: 806-749-5560
 Fax: 806-723-5272

Email: moliver@lbk.ars.usda.gov
 PCR primers
 FORWARD: GTTTTCCAGTCACGAC
 BACKWARD: CAGGAACAGCTATGAC.
 Location/Qualifiers
 1..569
 /organism="Tortula ruralis"
 /mol_type="mRNA"
 /db_xref="taxon:38588"
 /clone_id="Gametophyte rehydration Library"
 /note="Organ: Green Gametophyte; Vector: pSport1; Site_1: Salt; Site_2: NoCl"

ORIGIN

Query Match 4.6%; Score 38.2; DB 7; Length 569;
 Best Local Similarity 48.4%; Pred. No. 8.9;
 Matches 137; Conservative 0; Mismatches 143; Indels 3; Gaps 1;

QY 319 CTGCAAGAGCGCTTTGACCCGGATACCGCCTTATATGTCGCGTTGAAAGAGTTATG 378
 |||||
 Db 204 CTGGAGAGTCTTTTGGACAAGCCACAGATATTTCTTGCACGAAAGATAGTGTGTG 263
 |||||
 QY 379 CAGGTCGACCTGCGCCCTCGGGGTGGCGGATTACTGCGGGGCGCCTTTCGCTGTG 438
 |||||
 Db 264 GAGTTCAATGACGTGGAGACGGGACGAGCAATTCAGGGCTCTTCCACAGCAGAG 323
 |||||
 QY 439 GA--AAGCGAACACTGGGGGACGGCGGGCTATATTCCTCCGAAAAGCAGTCCGTTT 495
 |||||
 Db 324 GACTTGCGCGGGGCATATGTGGCGCGCGCCTCAGCCTGGCCGACATTCAAGCGCGGCC 383
 |||||
 QY 496 TTCCTGGAAGGTTTGGCGCCCTTCGCGCCCGGAAGGCTGCACCCGTGAATCTGATGATG 555
 |||||
 Db 384 ATGACGAGAGCGCGGTGACGCTACCCCGCAGTGACAGATGAAGAGGCGCATGAGAGATG 443
 |||||
 QY 556 TTCAGCGATTTTTCGACAGGGAAGAAATCCGTTTGCAGC 598
 |||||
 Db 444 TTCAGGCGCTGCAAGCAGAGGGGCTACATGATGAGCAGCATTC 486
 |||||

RESULT 14

LOCUS CF687694 553 bp mRNA linear EST 16-AUG-2004
 DEFINITION CCACV03TR C.neoformans strain JEC21 Cryptococcus neoformans var.
 neoformans cDNA clone CCACV03, mRNA sequence.
 ACCESSION CF687694
 VERSION CF687694.1 GI:41541853
 KEYWORDS EST.
 SOURCE Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
 var. neoformans)
 ORGANISM Cryptococcus neoformans var. neoformans
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
 Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
 Filobasidiella.
 1 (bases 1 to 553)
 Locus: B.
 End sequencing of clones from a full length enriched, normalized
 JEC21 cDNA library
 JOURNAL Unpublished (2003)
 COMMENT Other_ESTs: CCACV03TF
 Contact: Brendan Loftus
 The Institute for Genomic Research (TIGR; www.tigr.org)
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-3543
 Fax: 301-838-0208
 Email: b.loftus@tigr.org
 Seq primer: TR.
 Location/Qualifiers
 1..553
 /organism="Cryptococcus neoformans var. neoformans"
 /mol_type="mRNA"
 /strain="JEC21"
 /db_xref="taxon:40410"
 /clone="CCACV03"

FEATURES

source

Email: moliver@lbk.ars.usda.gov
 PCR primers
 FORWARD: GTTTTCCAGTCACGAC
 BACKWARD: CAGGAACAGCTATGAC.
 Location/Qualifiers
 1..569
 /organism="Tortula ruralis"
 /mol_type="mRNA"
 /db_xref="taxon:38588"
 /clone_id="Gametophyte rehydration Library"
 /note="Organ: Green Gametophyte; Vector: pSport1; Site_1: Salt; Site_2: NoCl"

ORIGIN

Query Match 4.6%; Score 38.2; DB 7; Length 569;
 Best Local Similarity 48.4%; Pred. No. 8.9;
 Matches 137; Conservative 0; Mismatches 143; Indels 3; Gaps 1;

QY 319 CTGCAAGAGCGCTTTGACCCGGATACCGCCTTATGCTCCGCTTGAAAGAGTTATG 378
 |||||
 Db 204 CTGGAGAGTCTTTTGGACAAGCCACAGATATTTCTTGCACGAAAGATGATGTTGTG 263
 |||||
 QY 379 CAGGTCGACCTGCGCCCTCGGGGTGGCGGATTACTGCGGGGCGCCTTTCGCTGTG 438
 |||||
 Db 264 GAGTTCAATGACGTGGAGACGGGAGACGGAGCAATTCAGGGCTCTCCACAGCAGAG 323
 |||||
 QY 439 GA--AAGCGAACACTGGGGGACGGCGGGCTATATTCATTTCCGAAAAGCANTCGGTTT 495
 |||||
 Db 324 GACTTGCGCGGGGCATAGTGTGGCGCGCGCCCTCAGCCTGGCCGACATTCAAGCGCGGCC 383
 |||||
 QY 496 TTCCTGGAAGGTTTGGCGCCCTTCGCGCCCGGAAGGCTGCACCCGTGAATCTGATGATG 555
 |||||
 Db 384 ATGACGAGAGCGCGGTGACGCTCACCCCGCAGTACAGATGAAGAGGCGCATGAGAAATG 443
 |||||
 QY 556 TTCAGCGATTTTTCGACAGGGAAGAAATCCGTTTGCAGC 598
 |||||
 Db 444 TTCAGGCGCTGCGACGACAGGGGCTACATGATGAGCAGCATTC 486
 |||||

RESULT 14

LOCUS CF687694 553 bp mRNA linear EST 16-AUG-2004
 DEFINITION CCACV03TR C.neoformans strain JEC21 Cryptococcus neoformans var.
 CF687694
 accession CFC687694
 version CF687694.1 GI:41541853
 keywords EST.
 SOURCE
 Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
 var. neoformans)
 ORGANISM
 Cryptococcus neoformans var. neoformans
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
 Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
 Filobasidiella.
 1 (bases 1 to 553)
 REFERENCE
 AUTHORS Loftus,B.
 TITLE End sequencing of clones from a full length enriched, normalized
 JEC21 cDNA library
 JOURNAL Unpublished (2003)
 COMMENT Other_ESTs: CCACV03TF
 Contact: Brendan Loftus
 The Institute for Genomic Research (TIGR; www.tigr.org)
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-3543
 Fax: 301-838-0208
 Email: b.loftus@tigr.org
 Seq primer: TR.
 FEATURES
 source Location/Qualifiers
 1..553
 /organism="Cryptococcus neoformans var. neoformans"
 /mol_type="mRNA"
 /strain="JEC21"
 /db_xref="taxon:40410"
 /clone="CCACV03"

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:17:47 ; Search time 296.232 Seconds
(without alignments)
4968.475 Million cell updates/sec

Title: US-09-211-691-1

Perfect score: 828

Sequence: 1 atgcaaacaccgcttaccag.....tcattgtgccttccaataa 828

Scoring table: IDENTITY_NIC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/PCTUS COMB.seq:*
- 7: /cgn2_6/prodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/prodata/1/ina/R COMB.seq:*
- 9: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	726	87.7	5859	2	US-08-312-387B-1
2	726	87.7	5859	2	US-08-312-387B-7
3	726	87.7	5859	2	US-08-683-426-1
4	726	87.7	5859	2	US-08-683-426-7
5	726	87.7	5859	2	US-08-683-458-1
6	726	87.7	5859	2	US-08-683-458-7
7	726	87.7	5859	2	US-08-878-360-1
8	726	87.7	5859	2	US-08-878-360-7
9	726	87.7	5859	3	US-08-878-140B-1
10	726	87.7	5859	3	US-09-333-412-1
11	726	87.7	5859	3	US-09-333-412-7
12	726	87.7	5859	3	US-09-338-943-1
13	726	87.7	5859	3	US-10-007-267A-1
14	726	87.7	5859	3	US-10-007-267A-7
15	86.6	10.5	1830121	3	US-09-557-884-1
16	86.6	10.5	1830121	3	US-09-643-990A-1
17	86.6	10.5	1830121	3	US-10-158-865-1
18	73.2	8.8	65792	3	US-09-540-236-487
19	73.2	8.8	65792	3	US-09-540-236-31
20	57	6.9	1830121	3	US-09-557-884-1
21	57	6.9	1830121	3	US-09-643-990A-1
22	57	6.9	1830121	3	US-10-158-865-1
23	39.4	4.8	4455	3	US-09-902-540-7119
24	39.4	4.8	4458	3	US-09-902-540-650

25	39	4.7	399	3	US-09-902-540-2937	Sequence 2937, App
26	39	4.7	23417	3	US-09-902-540-1207	Sequence 1207, App
27	38.6	4.7	753	3	US-09-252-991A-10290	Sequence 10290, A
28	38.6	4.7	1260	3	US-09-252-991A-10493	Sequence 10493, A
29	38.6	4.7	1635	3	US-09-252-991A-10567	Sequence 10567, A
30	38.6	4.7	1743	3	US-09-252-991A-10674	Sequence 10674, A
31	37.2	4.5	1401	4	US-09-605-703B-543	Sequence 543, App
32	36.8	4.4	31422	3	US-09-914-286-2	Sequence 2, App1
33	36	4.3	1721	2	US-07-688-352C-13	Sequence 13, App1
34	36	4.3	1721	2	US-08-474-375C-13	Sequence 13, App1
35	36	4.3	1721	3	US-09-146-249A-13	Sequence 13, App1
36	36	4.3	1721	3	US-08-206-188B-13	Sequence 13, App1
37	36	4.3	1721	6	PCT-US91-02714-13	Sequence 13, App1
38	35.8	4.3	1259	2	US-08-997-080-123	Sequence 123, App
39	35.8	4.3	1259	2	US-08-997-362-123	Sequence 123, App
40	35.8	4.3	1259	3	US-09-095-855-123	Sequence 123, App
41	35.8	4.3	1259	3	US-09-324-542-123	Sequence 123, App
42	35.8	4.3	1259	3	US-09-205-426-123	Sequence 123, App
43	35.8	4.3	2013	2	US-08-997-080-176	Sequence 176, App
44	35.8	4.3	2013	2	US-08-997-362-176	Sequence 176, App
45	35.8	4.3	2013	3	US-09-095-855-176	Sequence 176, App

ALIGNMENTS

RESULT 1
US-08-312-387B-1
; Sequence 1, Application US/08312387B
; Patent No. 5545553
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,387B
; FILING DATE: July 7, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELE: 133521
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5859 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Neisseria gonorrhoeae
; STRAIN: F62
; FEATURE:
; NAME/KEY: CDS

LOCATION: 1..381
FEATURE: CDS
NAME/KEY: 445..1491
LOCATION: 445..1491
FEATURE: CDS
NAME/KEY: 2342..3262
LOCATION: 2342..3262
FEATURE: CDS
NAME/KEY: 3322..4335
LOCATION: 3322..4335
FEATURE: CDS
NAME/KEY: 4354..5196
LOCATION: 4354..5196
US-08-312-387B-1

Query Match 87.7%; Score 726; DB 2; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2,8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

1 ATGCAAAACCAAGTATACAGCTTAGCTTCGCGGAGAAAGGCGGCGACATTCGCCGAT 60
1491 ATGCAAAACCAAGTATACAGCTTAGCTTCGCGGAGAAAGGCGGCGACATTCGCCGAT 1550
61 ACCTTCGCGAGGACGCGCATCCCGTTCAAGTTTTCGAGCGCATGATGCCGCTGAAAG 120
1551 ACCTTCGCGAGGACGCGCATCCCGTTCAAGTTTTCGAGCGCATGATGCCGCTGAAAG 1610
121 CTGGAACAGGCAATGCGGGAACCTCGTCCCGGCTTCGCGGAGAAAGGCGGCGACATTCGCCGAT 180
1611 CTGGAACAGGCAATGCGGGAACCTCGTCCCGGCTTCGCGGAGAAAGGCGGCGACATTCGCCGAT 1670
181 GTGGAAGAAAGCCCTCTTATAGAGCCAGCGCTATGTTGGAAGAGGAGCATTTGAGCGAAGT 240
1671 GTGGAAGAAAGCCCTCTTATAGAGCCAGCGCTATGTTGGAAGAGGAGCATTTGAGCGAAGT 1730
241 CTGCGCTATATACCGCTATTTGAGGACGAGCTTTACTCGCGGAGAGTGAAGAAATTC 300
1731 GTACCGTATATACCGCTATTTGAGGACGAGCTTTACTCGCGGAGAGTGAAGAAATTC 1790
301 CTGCGGAGAGGAGCTTGTGCTGCAAGAGCTTTGACCCGGAATACCGCTTTATTCGCGG 360
1791 CTGCGGAGAGGAGCTTGTGCTGCAAGAGCTTTGACCCGGAATACCGCTTTATTCGCGG 1850
361 TTGGAAGAGATGTTATAGCAAGCTCTGACCTCGCTCCGCGGTGCGGAGTATACGCGG 420
1851 TTGGAAGAGATGTTATAGCAAGCTCTGACCTCGCTCCGCGGTGCGGAGTATACGCGG 1910
421 CGGCGCTTCCGCTGTTGGAAGAGCACTGCGGAGAGCGCGGCTATATTCATTTCCGGA 480
1911 CGGCGCTTCCGCTGTTGGAAGAGCACTGCGGAGAGCGCGGCTATATTCATTTCCGGA 1970
481 AAAGCGAGTCCGCTTTTCTGTGACAGGTTGCGCGCTGCGCGGAGAGGCTGACCCG 540
1971 AAAGCGAGTCCGCTTTTCTGTGACAGGTTGCGCGCTGCGCGGAGAGGCTGACCCG 2030
541 GTGCGATGATATGTTGAGCGCATTTTTCGACAGGAGAAAGATGCGGCTTTCCAGCTC 600
2031 GTGCGATGATATGTTGAGCGCATTTTTCGACAGGAGAAAGATGCGGCTTTCCAGCTC 2090
601 AATCCGCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTACGAGCAAAAGCGCGA 660
2091 AATCCGCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTACGAGCAAAAGCGCGA 2150
661 TTGGGCAAGCTTATGAGACAGACGCGCTCTGAAACCGCAACAGCAAGGCGCATTC 720
2151 TTGGGCAAGCTTATGAGACAGACGCGCTCTGAAACCGCAACAGCAAGGCGCATTC 2210
721 CCGGCAACAGCATTTGAAACAGACGCGCTGATCGGCGCTTGAACCAATATGCGAGAAAG 780
2211 CCGGCAACAGCATTTGAAACAGACGCGCTGATCGGCGCTTGAACCAATATGCGAGAAAG 2270
781 GAAAAACGCGGCAAGGCGGCAAGCATTTGATG 814

Db 2271 GAAAAACGCGGCAAGGCGGCAAGCATTTGATG 2304

RESULT 2
US-08-312-387B-7
Sequence 7, Application US/08312387B
Patent No. 5545553

GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/312,387B
FILING DATE: July 7, 1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 5859 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Neisseria gonorrhoeae

STRAIN: F62

FEATURE:

NAME/KEY: CDS

LOCATION: 1491..2330

US-08-312-387B-7

Query Match 87.7%; Score 726; DB 2; Length 5859;

Best Local Similarity 93.2%; Pred. No. 2,8e-216;

Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

1 ATGCAAAACCAAGTATACAGCTTAGCTTCGCGGAGAAAGGCGGCGACATTCGCCGAT 60
1491 ATGCAAAACCAAGTATACAGCTTAGCTTCGCGGAGAAAGGCGGCGACATTCGCCGAT 1550
61 ACCTTCGCGAGGACGCGCATCCCGTTCAAGTTTTCGAGCGCATGATGCCGCTGAAAG 120
1551 ACCTTCGCGAGGACGCGCATCCCGTTCAAGTTTTCGAGCGCATGATGCCGCTGAAAG 1610
121 CTGGAACAGGCAATGCGGGAACCTCGTCCCGGCTTCGCGGAGAAAGGCGGCGACATTCGCCGAT 180
1611 CTGGAACAGGCAATGCGGGAACCTCGTCCCGGCTTCGCGGAGAAAGGCGGCGACATTCGCCGAT 1670
181 GTGGAAGAAAGCCCTCTTATAGAGCCAGCGCTATGTTGGAAGAGGAGCATTTGAGCGAAGT 240
1671 GTGGAAGAAAGCCCTCTTATAGAGCCAGCGCTATGTTGGAAGAGGAGCATTTGAGCGAAGT 1730

OY	241	CTGCCGATATACACGGTATTTTGGAGGACAGCGTTTACTCTGGCGGAAAGTGAAGAAAATTC	300
Db	1731	GTACCGTATATCCGCGTATTTGAAGATATATGCTTACTCTGGCGGAAAGGCGAGCAATTTC	1790
OY	301	CTTGGCCGAAGACGCTTGGCTCGCAAGAACGCTTTTGAACCCGGATATACCGCTTTATCGTCCGC	360
Db	1791	CTTGGCCGAAGATCTTGGCTCGCAAGAACGCTTTTGAACCCCGATTCGCGCTTTGTGTCGCGC	1850
OY	361	TTGGAAACGATGTTTATGCAAGTCTCTGACCTTCGCCCTTCGGCGGTGGCGGAATTACTCGCGGG	420
Db	1851	TTGGAAACGATGTTTATGCAAGTCTCTGACCTTCGCCCTTCGGCGGTGGCGGAATTACTCGCGGG	1910
OY	421	CGCGGCTTTCCGCTCTGTGGAAGAACGACATCTGGGGGAGGGCGGGCTATATATCTTTCCCGA	480
Db	1911	CGCGGCTTTCCGCTTTTGGAAAGCGCAACATCTGGGGAAGGGCGGGCTATATATTTTCCCGA	1970
OY	481	AAAGCGATGCGGATTTTCTCTGGAACAGTTTTGGCGGCTTCGCCCGCCGGAAGGGCTGCAACCC	540
Db	1971	AAAGCGATGCGGATTTTCTTGTGGAACAGTTTTGGCGGCTTCGCCCGCCGGAAGGGCTGCAACCTT	2030
OY	541	GTCGATCTGATGATGTTTCAAGCGATTTTTCGACAGGAAAGAAATGCGGTTTGCAGCTC	600
Db	2031	GTCGATTTGATGATGTTTTCGCGCAACCTTGACGACAGGAAAGAAATGCGGTTTGCAGCTC	2090
OY	601	AATCCCGCTTGTGTGGGCCCAAGAGCTGATTAATGCAAGTTTTCACGACCAAAACAGCGCA	660
Db	2091	AATCCCGCTTGTGTGGGCCCAAGAGCTGATTAATGCAAGTTTTCACGACCAAAACAGCGCA	2150
OY	661	TTGGGACAGCTGATGGAACAGACCGGCTCCTCGAAACGCGCAAAAGGCGGATTC	720
Db	2151	TTGGGACAGCTGATGGAACATGACCGCGCTGTAACGCGCAAAACAGCAATGCGGATTC	2210
OY	721	CCCGGCAACACATTCAAACACCGCTGTATCCGCGCTTTGACCAAAATCAGCAGGAAAAG	780
Db	2211	CCCGGCAACACATTCAAACACCGCTGTATCCGCGCTTTGACCAAAATGCGCAGGAAAAG	2270
OY	781	GAATAACGCGGCGCAAAAGCGCGCAACGTTCAATTG	814
Db	2271	GAATAACGCGGCGCAAAAGCGCGCAACGTTAATCG	2304
RESULT 3			
US-08-683-426-1			
; Sequence 1, Application US/08683426			
; Patent No. 5705367			
; GENERAL INFORMATION:			
; APPLICANT: Gotschlich, Emil C.			
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF			
; NUMBER OF SEQUENCES: 12			
; CORRESPONDENCE ADDRESSES:			
; ADDRESSEE: Klauber & Jackson			
; STREET: 411 Hackensack Avenue			
; CITY: Hackensack			
; STATE: New Jersey			
; COUNTRY: USA			
; ZIP: 07601			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/683,426			
; FILING DATE:			
; CLASSIFICATION: 536			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/312,387			
; FILING DATE: September 26, 1994			
; CLASSIFICATION: 536			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Jackson Esq., David A.			
; REGISTRATION NUMBER: 26,742			

Query Match	87.7%	Score 726	DB 2	Length 5859
Best Local Similarity	93.2%	Pred. No. 2.8e-216		
Matches 759	Conservative 0	Mismatches 55	Indels 0	Gaps 0
1	ATGCAAAACACGTTATCAGCTTACGCTTCCGCGCAGAAACGACAGGCGCATTTGCCGAT	60		
1491	ATGCAAAACACGTTATCAGCTTACGCTTCCGCGCAGAAACGACAGGCGCATTTGCCGAT	1550		
61	ACCTTCGCGACGACGCGCATCCGTTTCAGTTTTCGACGACGATGCGCTTGAAG	120		
1551	ACCTTCGCGACGACGCGCATCCGTTTCAGTTTTCGACGACGATGCGCTTGAAG	1610		
121	CTGGAACAGGAATGCGGGAATCTGTCGCCGCTTGTCCGCGCACCCCTTATTGAGGGA	180		
1611	CTGGAACAGGAATGCGGGAATCTGTCGCCGCTTGTCCGCGCACCCCTTATTGAGGGA	1670		
181	GTGAAAAAAGCTCTTATTAGACGACGCGTATTGTGAAGCAGGCAATTGAGCAAGGT	240		
1671	GTGAAAAAAGCTCTTATTAGACGACGCGTATTGTGAAGCAGGCAATTGAGCAAGGT	1730		
241	CTGCGGTATATCACCGTATTGAGACGACGTTTACTCGCGCAAGGTGAGAAAAATTC	300		
1731	GTACCGTATATCGCGGTATTGAGATGATCTTACTCGCGCAAGGTGAGAAAAATTC	1790		
301	CTTGCAGAAAGCGTTGCGTGCAGAAAGCTTGAACCGGATACCGGCTTATGTCGCG	360		
1791	CTTGCAGAAAGCTTGTGCTGCAGAAAGCTTGAACCGGATACCGGCTTATGTCGCG	1850		
361	TTGAAAAAGATGTTTATGACAGTCTGACCTCGCCCTCCGCGGTGAGATTAATGCGGG	420		
1851	TTGAAAAAGATGTTTATGACAGTCTGACCTCGCCCTCCGCGGTGAGATTAATGCGGG	1910		
421	CGCGCTTTTCGCTGTTGAAAAAGCACTGCGGAGCGCGCGCTTATCATTTCCGA	480		

Db 1911 CCGCGCTTTCCTTTTGGAAAGCACTGCGGGAACGGCGGCTATATTATTTCCGA 1970
Qy 481 AAAGCATGCGGTTTTTCTGGAAGATTTTGGCCCTGCGCCGCAAGGAGTGCACCC 540
Db 1971 AAGGCGATGCTTTTCTTCTGACAGGTTTGGCCGTTTGGCCGCAAGCCCTGACCC 2030
Qy 541 GTGATCTGATGATGTTGACGGATTTTTCAGAGGAAGATGCGGTTTCCAGCTC 600
Db 2031 GTGATTTGATGATGTTTGGCAACCTTGACAGACAGGAAGAAATGCGGTTTGCAGCTC 2090
Qy 601 AATCCGCTTGTGCGCCCAAGAGCTGATTCATTCAGAGTTTCAGACCAAAACAGCGCA 660
Db 2091 AATCCGCTTGTGCGCCCAAGAGCTGATTCATTCAGAGTTTCAGACCAAAACAGCGCA 2150
Qy 661 TTGGGAGCTGATTCGAAACGACGCGCTTCGACGCAAGCAAGAGGCGGATTC 720
Db 2151 TTGGGAGCTGATTCGAAACGACGCGCTTCGACGCAAGCAAGAGGCGGATTC 2210
Qy 721 CCGGCAACACATTCAAACCGCGCTGATTCGCGCTTGAACAAATCAGCAGGAAAG 780
Db 2211 CCGGCAACACATTCAAACCGCGCTGATTCGCGCTTGAACAAATCAGCAGGAAAG 2270
Qy 781 GAAAAAGCGCGGCAAGCGCGCAAGCTTCATTG 814
Db 2271 GAAAAAGCGCGGCAAGCGCGCAAGCTTCATTG 2304

RESULT 4

US-08-683-426-7
Sequence 7, Application US/08683426

Patent No. 5705367

GENERAL INFORMATION:

APPLICANT: GoeSchlich, Emil C.

TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF

TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSER: Klauder & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/683,426

FILING DATE:

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/312,387

FILING DATE: September 26, 1994

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-0958

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 5859 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:
ORGANISM: Neisseria gonorrhoeae
STRAIN: F62
FEATURE:
NAME/KEY: CDS
LOCATION: 1491..2330
OTHER INFORMATION: 19cB
US-08-683-426-7

Query Match 87.7%; Score 726; DB 2; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 1 ATGCAAAACACGTTATGACCTTACCTCCGCGGAGAAACGAGGCGGACATTCGCGAT 60
Db 1491 ATGCAAAACACGTTATGACCTTACCTCCGCGGAGAAACGAGGCGGACATTCGCGCA 1550
Qy 61 ACCCTCGGACAGGACGAGCATCCGTTTGAATTTTCAGACGATGATCCGCTGAAAG 120
Db 1551 ACCCTCGGACAGGACGAGCATCCGTTTGAATTTTCAGACGATGATCCGCTGAAAG 1610
Qy 121 CTGGAACGGGCAATGCGGAACTGCTCCCGGCTTGTGCGGCAACCCCTATTTGAGCGGA 180
Db 1611 CTGGAACGGGCAATGCGGAACTGCTCCCGGCTTGTGCGGCAACCCCTATTTGAGCGGA 1670
Qy 181 GTGAAAAAGCTGCTTTATGAGCCAGCGCTTATTTGTGAAAGCAGGCAATTGACAGAGT 240
Db 1671 GTGAAAAAGCTGCTTTATGAGCCAGCGCTTATTTGTGAAAGCAGGCAATTGACAGAGC 1730
Qy 241 CTGCGCTATATCACCGTATTTGAGGACGACGTTTACTCGGGAAGGATGAGAAATTC 300
Db 1731 GTACGTTATATCCCGTATTTGAGGATGATCTTACTCGGGAAGGCGGAGATTC 1790
Qy 301 CTGCGCAAGACGCTTGGCTGCAAGACGCTTTGACCCGGAATCCGCTTATGCTCCG 360
Db 1791 CTGCGCAAGATCTTGGCTGCAAGACGCTTTGACCCGGAATCCGCTTATGCTCCG 1850
Qy 361 TTGAAACGATGTTATGACGCTCTGACCTCGCCTCGGCGGAGGAGATTAATGCGGAG 420
Db 1851 TTGAAACGATGTTATGACGCTCTGACCTCGCCTCGGCGGAGGAGATTAATGCGGAG 1910
Qy 421 CCGGCTTTCCGCTGTGAAAGGCAACCTGGGGGAGCGCGGCTATATCATTTCCCGA 480
Db 1911 CCGGCTTTCCGCTGTGAAAGGCAACCTGGGGGAGCGCGGCTATATTAATTTCCCGA 1970
Qy 481 AAAGCATGCGGTTTTTCTGGAAGATTTTGGCCCTGCGCCGCAAGGAGTGCACCC 540
Db 1971 AAAGCATGCGGTTTTTCTGGAAGATTTTGGCCCTGCGCCGCAAGGAGTGCACCC 2030
Qy 541 GTGATCTGATGATGTTGACGGATTTTTCAGAGGAAGATGCGGTTTCCAGCTC 600
Db 2031 GTGATTTGATGATGTTTGGCAACCTTGACAGACAGGAAGAAATGCGGTTTGCAGCTC 2090
Qy 601 AATCCGCTTGTGCGCCCAAGAGCTGATTCATTCAGAGTTTCAGACCAAAACAGCGCA 660
Db 2091 AATCCGCTTGTGCGCCCAAGAGCTGATTCATTCAGAGTTTCAGACCAAAACAGCGCA 2150
Qy 661 TTGGGAGCTGATTCGAAACGACGCGCTTCGACGCAAGCAAGAGGCGGATTC 720
Db 2151 TTGGGAGCTGATTCGAAACGACGCGCTTCGACGCAAGCAAGAGGCGGATTC 2210
Qy 721 CCGGCAACACATTCAAACCGCGCTGATTCGCGCTTGAACAAATCAGCAGGAAAG 780
Db 2211 CCGGCAACACATTCAAACCGCGCTGATTCGCGCTTGAACAAATCAGCAGGAAAG 2270
Qy 781 GAAAAAGCGCGGCAAGCGCGCAAGCTTCATTG 814
Db 2271 GAAAAAGCGCGGCAAGCGCGCAAGCTTCATTG 2304

RESULT 5

US-08-683-458-1

Sequence 1, Application US/08683458

Patent No. 5798233
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,458
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5859 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULAR TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Neisseria gonorrhoeae
STRAIN: F62
FEATURE:
NAME/KEY: CDS
LOCATION: 1..381
OTHER INFORMATION: glys (glycyl tRNA synthetase beta chain)
FEATURE:
NAME/KEY: CDS
LOCATION: 445..1491
OTHER INFORMATION: 1gfa
FEATURE:
NAME/KEY: CDS
LOCATION: 2342..3262
OTHER INFORMATION: 1gfc
FEATURE:
NAME/KEY: CDS
LOCATION: 3322..4335
OTHER INFORMATION: 1gfd
FEATURE:
NAME/KEY: CDS
LOCATION: 4354..5196
OTHER INFORMATION: 1gre
US-08-683-458-1

Query Match 87.7%; Score 726; DB 2; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGCAAAACCACTTATCAGCTTAGCTCCGCGCGAAGCGAGGCGCAGATTGGCGAT 60

Db 1491 ATGCAAAACCACTTATCAGCTTAGCTCCGCGCGAAGCGAGGCGCAGATTGGCGCA 1550
QY 61 ACCTTGCGAGGACGCGCATCCCGTTTCAGTTTTCAGCGACATGATCCGTGTGAAGG 120
Db 1551 ACCTTGCGAGGATGCGCGCATCCCGTTTCAGTTTTCAGCGACATGATCCGTGTGAAGG 1610
QY 121 CTGGAACAGGCATGCGGGAACCGCCCGCTTCGCGGACCCCTATTGACCGGA 180
Db 1611 CTGGAACAGGCATGCGGGAACCGCCCGCTTCGCGGACCCCTATTGACCGGA 1670
QY 181 GTGAAAAAGCCCTGTTTATGAGCCAGCGCGTATTGTGGAAGGACATTGACGAAGT 240
Db 1671 GTGAAAAAGCCCTGTTTATGAGCCAGCGCGTATTGTGGAAGGACATTGACGAAGG 1730
QY 241 CTGCGGTATATCAGCGTATTGAGAGCAGTTTATCTCGCGAAGTGAAGAAAAATTC 300
Db 1731 GTACCGTATATGCGCGTATTGAGAGTATGTTCTTCTCGGCGAAGGCGGAGCAGTTTC 1790
QY 301 CTGCGCGAAGACGCTTGCTGCAAGACGCTTGAACCGCGATACCGCTTATGTCGCGC 360
Db 1791 CTGCGCGAAGATACCTTGCTGCAAGACGCTTGAACCGCGATACCGCTTATGTCGCGC 1850
QY 361 TTGAAAACGATGTTTATGACAGTCTGACGCTCGCGCGTGGCGGATTAATGCGGG 420
Db 1851 TTGAAAACGATGTTTATGACAGTCTGACGCTCGCGCGTGGCGGATTAATGCGGG 1910
QY 421 CGGCGCTTTCGCGTGTGGAAGCGAACACTGCGGAGCGCGGCTATATCAATTTCCGA 480
Db 1911 CGGCGCTTTCGCGTGTGGAAGCGAACACTGCGGAGCGCGGCTATATCAATTTCCGA 1970
QY 481 AAGGCGATCGGTTTCTGTCGACAGGTTTGGCCCTCGCGCGGAAGGCTGACCCC 540
Db 1971 AAGGCGATCGGTTTCTGTCGACAGGTTTGGCCCTCGCGCGGAAGGCTGACCCC 2030
QY 541 GTGCAATCTGATGATGTTGACGATTTTTCGACAGGAGGAATGCGGTTTCCAGCTC 600
Db 2031 GTGCAATCTGATGATGTTGACGATTTTTCGACAGGAGGAATGCGGTTTCCAGCTC 2090
QY 601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCCAATTTTCAGACCCAAACAGCGCA 660
Db 2091 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCCAATTTTCAGACCCAAACAGCGCA 2150
QY 661 TTGGGCGAGCTTATGGAACAGACCGCTTCTGAAACCGGAACAGGAAGGCGCATTC 720
Db 2151 TTGGGCGAGCTTATGGAACAGACCGCTTCTGAAACCGGAACAGGAAGGCGCATTC 2210
QY 721 CCCGCAACACATTCAAACACCGCGTATCCGCGCTTGACCAAAATGACGAGGAAAGG 780
Db 2211 CCCGCAACACATTCAAACACCGCGTATCCGCGCTTGACCAAAATGACGAGGAAAGG 2270
QY 781 GAAAAACGCGCGCAAAAGGCGGACAGATTCAATG 814
Db 2271 GAAAAACGCGCGCAAAAGGCGGACAGATTCAATG 2304

RESULT 6
US-08-683-458-7
Sequence 7, Application US/08683458
Patent No. 5798233
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:

```

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,458
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5859 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Neisseria gonorrhoeae
STRAIN: F62
FEATURE:
NAME/KEY: CDS
LOCATION: 1491..2330
OTHER INFORMATION: 15TB
US-08-683-458-7

Query Match      87.7%; Score 726; DB 2; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2,8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGCAAAACGCGTATGCTTACGCTTCCGCGCAAGCGAGGCGGCACTTGGCCAT 60
DB 1491 ATGCAAAACGCGTATGCTTACGCTTCCGCGCAAGCGAGGCGGCACTTGGCCGA 1550
QY 61 ACCCTCGGAGGCGACGCGCATCCCGTTTCAGTTTTCGACGACGCTGCTGAAAG 120
DB 1551 ACCCTCGGAGTGGGGGATCCCGTTTCAGTTTTCGACGACGCTGCTGAAAG 1610
QY 121 CTGGAACAGGCAATGGCGCACTCGTCCCGGCTGTGCGGCGACCCCTATTGAGCGGA 180
DB 1611 CTGGAACAGGCAATGGCGCACTCGTCCCGGCTGTGCGGCGACCCCTATTGAGCGGA 1670
QY 181 GTGGAAGAAAGCTGCTTATGAGCAGCGCGCTATTGTGAGACAGGCAATGGAAGAGGT 240
DB 1671 GTGGAAGAAAGCTGCTTATGAGCAGCGCGCTATTGTGAGACAGGCAATGGAAGAGGT 1730
QY 241 CTGCGGTATATACCGTATTGAGACGAGCTTTTACTCGGCGAAAGGTGAGAAATATTC 300
DB 1731 GTACCGTATATCGCGTATTGAGACGAGCTTTTACTCGGCGAAAGGTGAGAAATATTC 1790
QY 301 CTTGCGGAAGAGCTTGGCTGCAAGAGCTTGAACCGGGAATACCGCTTTATCTCGGC 360
DB 1791 CTTGCGGAAGAGCTTGGCTGCAAGAGCTTGAACCGGGAATACCGCTTTATCTCGGC 1850
QY 361 TTGGAAGAGATTTATGACGCTCTGACCTTCCGCGCTGCGGCTGAGGAGATTACTGCGGG 420
DB 1851 TTGGAAGAGATTTATGACGCTCTGACCTTCCGCGCTGCGGCTGAGGAGATTACTGCGGG 1910
QY 421 CGCGGCTTTCGCGCTGTGGAAGCGAAGCTGCGGGAAGCGGCGCTATATCTATTTCCCGGA 480
DB 1911 CGCGGCTTTCGCGCTTTTGAAGCGAAGCTGCGGGAAGCGGCGCTATATTTTCCCGGA 1970
```

```

QY 481 AAAGCATCGGCTTTTCTGACAGAGTTTCCGCGCTGCGCGCGGAAGGCGTCGACCC 540
DB 1971 AAAGCATCGGCTTTTCTGACAGAGTTTCCGCGCTGCGCGCGGAAGGCGTCGACCC 2030
QY 541 GTGCATCTGATGATGTTGAGGATTTTTCACAGGGAAGGAATGCGGTTTGCAGCTC 600
DB 2031 GTGCATTTGATGATGTTGAGGATTTTTCACAGGGAAGGAATGCGGTTTGCAGCTC 2090
QY 601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTTATGCAAGTTTCAAGCAAAAGCGCA 660
DB 2091 AATCCGCGCTTGTGCGCCCAAGAGCTGATTTATGCAAGTTTCAAGCAAAAGCGCA 2150
QY 661 TTGGGACACCTGATGAAACAGACCGCTCTGAAACCGCAACGCAAGGCGGATTCC 720
DB 2151 TTGGGACACCTGATGAAACAGACCGCTCTGAAACCGCAACGCAATGCGGATTCC 2210
QY 721 CCGGCAACATTTCAACACCGCTGATCCGCGCTTGAACCAATCAGCAGGGAAGG 780
DB 2211 CCGGCAACATTTCAACACCGCTGATCCGCGCTTGAACCAATCAGCAGGGAAGG 2270
QY 781 GAAAAACGCGCGCAAGGCGCGAAACAGTTCAATTG 814
DB 2271 GAAAAACGCGCGCAAGGCGCGAAACAGTTAATCG 2304

RESULT 7
US-08-878-360-1
Sequence 1, Application US/08878360
Patent No. 5945322
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauder & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,360
FILING DATE: 18-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/683,426
FILING DATE:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5859 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
```

ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Neisseria gonorrhoeae
STRAIN: F62
FEATURE:
NAME/KEY: CDS
LOCATION: 1..381
OTHER INFORMATION: glys (glycyl tRNA synthetase beta chain)
FEATURE:
NAME/KEY: CDS
LOCATION: 445..1491
OTHER INFORMATION: 1gca
FEATURE:
NAME/KEY: CDS
LOCATION: 2342..3262
OTHER INFORMATION: 1gcb
FEATURE:
NAME/KEY: CDS
LOCATION: 3322..4335
OTHER INFORMATION: 1gcb
FEATURE:
NAME/KEY: CDS
LOCATION: 4354..5196
OTHER INFORMATION: 1gcb
US-08-878-360-1

Query Match 87.7%; Score 726; DB 2; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.8e-216;

Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGCAAAACCAAGTATACAGCTTACGCTTCCGCGCAAGACGCGGCGCATTTGCCGAT 60
DB 1491 ATGCAAAACCAAGTATACAGCTTCCGCGCAAGACGCGGCGCATTTGCCGCA 1550
QY 61 ACCTTGGCGAGGACGCGCATCCGCTTCACTTTTGGAGCGCATGATCCGCTGAAAG 120
DB 1551 ACCTTGGCGAGTGCAGCATCCGCTTCACTTTTGGAGCGCATGATCCGCTGAAAG 1610
QY 121 CTGGAACAGGCAATGGCGGAACCTGCTCCGCGCTGTCGCGCGGACCCCTATTGAGCGGA 180
DB 1611 CTGGAACAGGCAATGGCGGAACCTGCTCCGCGCTGTCGCGCGGACCCCTATTGAGCGGA 1670
QY 181 GTGAAAAAGCCTGCTTATGAGCCAGCGCTATGTTGGAAGAGGCAATTGAGCAAGGT 240
DB 1671 GTGAAAAAGCCTGCTTATGAGCCAGCGCTATGTTGGAAGAGGCAATTGAGCAAGGT 1730
QY 241 CTGCGCTATATCAACGCTATTGAGAGCAAGTTTATCTCGGCGAAGGTGAGAAAAATTC 300
DB 1731 GTACCGTATATCGCCGCTATTGAGAGCAAGTTTATCTCGGCGAAGGTGAGAAAAATTC 1790
QY 301 CTGCGCGAAGACGCTTGGCTGCAAGAACGCTTGGACCGGATACCGCTTATGTCGCG 360
DB 1791 CTGCGCGAAGATACGCTTGGCTGCAAGAACGCTTGGACCGGATACCGCTTATGTCGCG 1850
QY 361 TTGGAACAGATGTTATGACGCTGACCTGCGCTCGCGGCGTGGCGATTAATGCGGG 420
DB 1851 TTGGAACAGATGTTATGACGCTGACCTGCGCTCGCGGCGTGGCGATTAATGCGGG 1910
QY 421 CGCGCTTTCGCTGTTGGAAGAGCAACCTGCGGAGCGCGGCTATATCAATTTCCGA 480
DB 1911 CGCGCTTTCGCTGTTGGAAGAGCAACCTGCGGAGCGCGGCTATATCAATTTCCGA 1970
QY 481 AAAGGAGATGCGGTTTTCCGTCAGAGGTTGCGCGCTGCGCGCGGAGGCTGACCCG 540
DB 1971 AAAGGAGATGCGGTTTTCCGTCAGAGGTTGCGCGCTGCGCGCGGAGGCTGACCCG 2030
QY 541 GTGATCTGATGATGTTCAAGCATTTTTCGACAGGAGAAATGCGGTTTGCCAGCTC 600
DB 2031 GTGATCTGATGATGTTGCGCAACCTGACAGAGGAGAAATGCGGTTTGCCAGCTC 2090
QY 601 AATCCGCTTGTGCGCGCAAGAGTGCATTAATGCAAGTTTCAAGCAAAACAGCGCA 660
DB 2091 AATCCGCTTGTGCGCGCAAGAGTGCATTAATGCAAGTTTCAAGCAAAACAGCGCA 2150

QY 661 TTGAGAGCGCTGATGACACGACGCGCTTCTGTAACCGCAAGACAGGCGCGATTC 720
DB 2151 TTGAGAGCGCTGATGACACGACGCGCTTCTGTAACCGCAAGACAGGCGCGATTC 2210
QY 721 CCGCGCAACATTCATAAGACGCGCTGATCCGCGCTTGAACCAAAATGACGAGGAAG 780
DB 2211 CCGCGCAACATTCATAAGACGCGCTGATCCGCGCTTGAACCAAAATGACGAGGAAG 2270
QY 781 GAAAAAGCGCGCAAGAGCGCGCAACGTTGATG 814
DB 2271 GAAAAAGCGCGCAAGAGCGCGCAACGTTAATG 2304

RESULT 8

US-08-878-360-7
Sequence 7, Application US/08878360
Patent No. 5945322
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,360
FILING DATE: 18-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/683,426
FILING DATE:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5859 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Neisseria gonorrhoeae
STRAIN: F62
FEATURE:
NAME/KEY: CDS
LOCATION: 1491..2330
OTHER INFORMATION: 1gcb
US-08-878-360-7

Query Match 87.7%; Score 726; DB 2; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.8e-216;

```
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 1 ATGCAAAACCACTTATCAGCTTAGCTTCCGCGCAAGACGCGGCGCATTTGCCGAT 60
DB 1491 ATGCAAAACCACTTATCAGCTTAGCTTCCGCGCAAGACGCGGCGCATTTGCCGCA 1550
QY 61 ACCCTGCGGAGGACGCGCATCCCGTTTCAGTTTTTTCAGCATCTGACGCTGTAAG 120
DB 1551 ACCCTGCGGAGTGCAGCATCCCGTTTCAGTTTTTTCAGCATCTGACGCTGTAAG 1610
QY 121 CTGGAACAGGCAATGCGGCACTGCGTCCCGGCTTGTCCGCGCAACCCCTATTGAGCGGA 180
DB 1611 CTGGAACAGGCAATGCGGCACTGCGTCCCGGCTTGTCCGCGCAACCCCTATTGAGCGGA 1670
QY 181 GTGGAAGAAAGCGCTGCTTTATGAGCCAGCGCGTATTGTGGAAGAGGCAATTGACGAAGGT 240
DB 1671 GTGGAAGAAAGCGCTGCTTTATGAGCCAGCGCGTATTGTGGAAGAGGCAATTGACGAAGGT 1730
QY 241 CTGCGGTATATCACCCTATTGAGAGCAGCGTTTATCTCGGCGAAGGTGAGAAAAATTC 300
DB 1731 GTACCGTATATCACCCTATTGAGAGCAGCGTTTATCTCGGCGAAGGTGAGAAAAATTC 1790
QY 301 CTGCGCAAGAGCGCTTGGCTGCAAGAACGCTTTGACCGGATACCGGCTTATCTGCGCG 360
DB 1791 CTGCGCAAGATCTTGGCTGCAAGAACGCTTTGACCGGATACCGGCTTATCTGCGCG 1850
QY 361 TTGGAACAGATGTTTATGACAGCTCTGACCTCGCCCTCGCGGCGAGTATTACTGCGG 420
DB 1851 TTGGAACAGATGTTTATGACAGCTCTGACCTCGCCCTCGCGGCGAGTATTACTGCGG 1910
QY 421 CCGGCGCTTCCGCTGTTTGAAGAACGCACTGGGGGAGCGCGGCTATATCATTTCCGA 480
DB 1911 CCGGCGCTTCCGCTGTTTGAAGAACGCACTGGGGGAGCGCGGCTATATTTTCCGA 1970
QY 481 AAAGCGATCGGTTTTTCTGACAGGTTTGCCTCGCCGCGCAAGGCGTGACCCG 540
DB 1971 AAAGCGATCGGTTTTTCTGACAGGTTTGCCTCGCCGCGCAAGGCGTGACCCG 2030
QY 541 GTGATCTGATGATGTTTCAGCGATTTTTTTCAGACGGAAGAAATGCGGTTTGCAGCTC 600
DB 2031 GTGATCTGATGATGTTTCAGCGATTTTTTTCAGACGGAAGAAATGCGGTTTGCAGCTC 2090
QY 601 AATCCCGCTTGTGCGCCCAAGAGTGCATTAATGCAAGTTTTCAGACGCAAAACGCGCA 660
DB 2091 AATCCCGCTTGTGCGCCCAAGAGTGCATTAATGCAAGTTTTCAGACGCAAAACGCGCA 2150
QY 661 TTGGGCGAGCTGATGCAACAGCGCTCTCTGAACCGCAAAACGCAAAAGCGCAATTC 720
DB 2151 TTGGGCGAGCTGATGCAACAGCGCTCTCTGAACCGCAAAACGCAATGCGCGCATTC 2210
QY 721 CCGGCAACACATTTCAAAACAGCGCGCTGATCCGCGCTTGAACCAAAATCGACGGGAAAG 780
DB 2211 CCGGCAACACATTTCAAAACAGCGCGCTGATCCGCGCTTGAACCAAAATCGGCAAGG 2270
QY 781 GAAAAAGCGCGCAAGGCGCGAAGCATTTG 814
DB 2271 GAAAAAGCGCGCAAGGCGCGAAGCATTTG 2304

RESULT 9
US-08-478-140B-1
/ Sequence 1, Application US/08478140B
/ Patent No. 6127153
/ GENERAL INFORMATION:
/ APPLICANT: JOHNSON, KARL F.
/ APPLICANT: BOTH, STEPHEN
/ APPLICANT: BUTZALA, STEPHANIE L.
/ TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO
/ TITLE OF INVENTION: SACCHARIDE UNITS WITH A POLYGLYCOSYLTRANSFERASE, A
/ TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE AND GENE ENCODING A
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
```

```
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,140B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Laura A. Coruzzi
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7188-017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5859 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
ANTI-SENSE: NO
ORIGINAL SOURCE:
STRAIN: F62
FEATURE:
NAME/KEY: CDS
LOCATION: 1..381
FEATURE:
NAME/KEY: CDS
LOCATION: 445..1491
FEATURE:
NAME/KEY: CDS
LOCATION: 2342..3262
FEATURE:
NAME/KEY: CDS
LOCATION: 3322..4335
FEATURE:
NAME/KEY: CDS
LOCATION: 4354..5196
US-08-478-140B-1

Query Match 87.7%; Score 726; DB 3; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 1 ATGCAAAACCACTTATCAGCTTAGCTTCCGCGCAAGACGCGGCGCATTTGCCGAT 60
DB 1491 ATGCAAAACCACTTATCAGCTTAGCTTCCGCGCAAGACGCGGCGCATTTGCCGCA 1550
QY 61 ACCCTGCGGAGGACGCGCATCCCGTTTCAGTTTTTTCAGCATCTGACGCTGTAAG 120
DB 1551 ACCCTGCGGAGTGCAGCATCCCGTTTCAGTTTTTTCAGCATCTGACGCTGTAAG 1610
QY 121 CTGGAACAGGCAATGCGGCACTGCGTCCCGGCTTGTCCGCGCAACCCCTATTGAGCGGA 180
DB 1611 CTGGAACAGGCAATGCGGCACTGCGTCCCGGCTTGTCCGCGCAACCCCTATTGAGCGGA 1670
QY 181 GTGGAAGAAAGCGCTGCTTTATGAGCCAGCGCGTATTGTGGAAGAGGCAATTGACGAAGGT 240
DB 1671 GTGGAAGAAAGCGCTGCTTTATGAGCCAGCGCGTATTGTGGAAGAGGCAATTGACGAAGGT 1730
QY 241 CTGCGGTATATCACCCTATTGAGAGCAGCGTTTATCTCGGCGAAGGTGAGAAAAATTC 300
DB 1731 GTACCGTATATCACCCTATTGAGAGCAGCGTTTATCTCGGCGAAGGTGAGAAAAATTC 1790
```

QY 301 CTGCGGAGACGCTGGCTGCAAGAACGCTTTGACCCGGATACCGCCTTTATGTCGCG 360
DB 1791 CTGCGGAGATCTTGCTGCAAGAACGCTTTGACCCGGATACCGCCTTTATGTCGCG 1850
QY 361 TTGGAACGATGTTTATGACAGCTCTGACCTCGCCTCGGCGTGGGGAATTAATGCGGG 420
DB 1851 TTGGAACGATGTTTATGACAGCTCTGACCTCGCCTCGGCGTGGGGAATTAATGCGGG 1910
QY 421 CGGCGCTTTCGCGTTTGGAAGAGAACACTGGGGGACCGCGGCTTATCATTTTCCCGA 480
DB 1911 CGGCGCTTTCGCGTTTGGAAGAGAACACTGGGGGACCGCGGCTTATCATTTTCCCGA 1970
QY 481 AAGGCAATGCGTTTCTCTGCAAGGTTGCGCGCTCGCGCCGCAAGGCGTGCACCC 540
DB 1971 AAGGCAATGCGTTTCTCTGCAAGGTTGCGCGCTCGCGCCGCAAGGCGTGCACCC 2030
QY 541 GTGATCTGATGATGTTGACGCAATTTTTCAGACAGGAAATGCGGTTTGCCAGCTC 600
DB 2031 GTGATCTGATGATGTTGACGCAATTTTTCAGACAGGAAATGCGGTTTGCCAGCTC 2090
QY 601 AATCCGCGCTTGCGCCGCAAGAGCTGCAATTAATGCAAGTTTCAACCAACAGCGCA 660
DB 2091 AATCCGCGCTTGCGCCGCAAGAGCTGCAATTAATGCAAGTTTCAACCAACAGCGCA 2150
QY 661 TTGGGAGAGCTGATGCAACAGACCGCTCTGCAACGCAAAAGGCGGCAATTC 720
DB 2151 TTGGGAGAGCTGATGCAACAGACCGCTCTGCAACGCAAAAGGCGGCAATTC 2210
QY 721 CCGGCAACATTCATAACACCGCTGATCGCGCTTGAACCAAAATCAGCAGGAAAG 780
DB 2211 CCGGCAACATTCATAACACCGCTGATCGCGCTTGAACCAAAATCAGCAGGAAAG 2270
QY 781 GAAAAACGCGGCAAGGCGGCAACGTTCAATG 814
DB 2271 GAAAAACGCGGCAAGGCGGCAACGTTCAATG 2304

RESULT 10
US-09-333-412-1

GENERAL INFORMATION:
APPLICANT: Gottschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSES: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/333,412
FILING DATE: 15-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: July 7, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-333-412-1

Query Match 87.7%; Score 726; DB 3; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2,8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGCAAAACACGTTATGAGCTTAGCTTCGCGCGAAGAACGAGGCGGACATTTGCGAT 60
DB 1491 ATGCAAAACACGTTATGAGCTTAGCTTCGCGCGAAGAACGAGGCGGACATTTGCGCG 1550
QY 61 ACCTTGAGGACGACGCGCATCCCGTTTTCAGTTTTCAGACGCACTGATGCGTCTGAAG 120
DB 1551 ACCTTGAGGACGACGCGCATCCCGTTTTCAGTTTTCAGACGCACTGATGCGTCTGAAG 1610
QY 121 CTGGAACAGGCAATGCGGAACTCGTCCCGCTTGTGCGCGCACCTTAATTTGACGGA 180
DB 1611 CTGGAACAGGCAATGCGGAACTCGTCCCGCTTGTGCGCGCACCTTAATTTGACGGA 1670
QY 181 GTGAAAAAAGCGCTTATGAGCCAGCGCGTATGTTGGAAGAGCATTTGACGAAGGT 240
DB 1671 GTGAAAAAAGCGCTTATGAGCCAGCGCGTATGTTGGAAGAGCATTTGACGAAGGT 1730
QY 241 CTGCGTATATGACCGTATTTGAGGACGATTTTACTGCGGAAAGTGAAGAAATTC 300
DB 1731 GTACGATATATGCGCGTATTTGAAATGATGTTTACTGCGGAAAGTGAAGAGAGCTTC 1790
QY 301 CTGCGGAAAGACGCTTGCTGCAAGACGCTTTGACCGGAATACCGCTTTATGCTCGC 360
DB 1791 CTGCGGAAAGATATGCTGCTGCAAGACGCTTTGACCGGAATACCGCTTTATGCTCGC 1850
QY 361 TTGGAAGAGATGTTATGACGCTGCAAGCTCGCTCGCGGCTGCGGATTAATGCGGG 420
DB 1851 TTGGAAGAGATGTTATGACGCTGCAAGCTCGCTCGCGGCTGCGGATTAATGCGGG 1910
QY 421 CCGCGCTTTCGCTGTTGGAAGGCAACACTGGGGAAGCGGCGTATATATTCGCGA 480
DB 1911 CCGCGCTTTCGCTGTTGGAAGGCAACACTGGGGAAGCGGCGTATATATTCGCGA 1970
QY 481 AAGGCAATGCGTTTCTCTGCAAGGTTGCGCGCTCGCGCCGCAAGGCGTGCACCC 540
DB 1971 AAGGCAATGCGTTTCTCTGCAAGGTTGCGCGCTCGCGCCGCAAGGCGTGCACCC 2030
QY 541 GTGATCTGATGATGTTGACGCAATTTTTCAGACAGGAAATGCGGTTTGCCAGCTC 600
DB 2031 GTGATCTGATGATGTTGACGCAATTTTTCAGACAGGAAATGCGGTTTGCCAGCTC 2090
QY 601 AATCCGCGCTTGCGCCGCAAGAGCTGCAATTAATGCAAGTTTCAACCAACAGCGCA 660
DB 2091 AATCCGCGCTTGCGCCGCAAGAGCTGCAATTAATGCAAGTTTCAACCAACAGCGCA 2150
QY 661 TTGGGAGAGCTGATGCAACAGACCGCTCTGCAACGCAAAAGGCGGCAATTC 720
DB 2151 TTGGGAGAGCTGATGCAACAGACCGCTCTGCAACGCAAAAGGCGGCAATTC 2210
QY 721 CCGGCAACATTCATAACACCGCTGATCGCGCTTGAACCAAAATCAGCAGGAAAG 780
DB 2211 CCGGCAACATTCATAACACCGCTGATCGCGCTTGAACCAAAATCAGCAGGAAAG 2270
QY 781 GAAAAACGCGGCAAGGCGGCAACGTTCAATG 814
DB 2271 GAAAAACGCGGCAAGGCGGCAACGTTCAATG 2304

RESULT 11
US-09-333-412-7

GENERAL INFORMATION:
APPLICANT: Gottschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSES: Klauber & Jackson
STREET: 411 Hackensack Avenue

CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/09/333,412
APPLICATION NUMBER: US/09/333,412
FILING DATE: 15-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: July 7, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Bsg., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-333-412-7
Query Match 87.7%; Score 726; DB 3; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2,8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
1 ATGCAAAACCAAGTTATGAGCTTAGCTCCGCGAGAGGAGGCGACATTCGCCAT 60
1491 ATGCAAAACCAAGTTATGAGCTTAGCTCCGCGAGAGGAGGCGACATTCGCCA 1550
61 ACCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
1551 ACCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1610
121 CTGGAACAGGCAATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
1611 CTGGAACAGGCAATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1670
181 GTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
1671 GTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1730
241 CTGCGGTATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
1731 GTACGGTATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1790
301 CTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
1791 CTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1850
361 TTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
1851 TTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1910
421 GCGGCGCTTCCGCTGTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
1911 GCGGCGCTTCCGCTGTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1970
481 AAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
1971 AAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2030
541 GTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
2031 GTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2090
601 AATCCGCTTGTGCGGCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660

2091 AATCCGCTTGTGCGGCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2150
661 TTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
2151 TTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2210
721 CCGGCAACATTCATTAACACCGGCTGATCCGCGCTTGACCAAAATGACGAGGAGG 780
2211 CCGGCAACATTCATTAACACCGGCTGATCCGCGCTTGACCAAAATGAGGAGGAGG 2270
781 GAAAAACCGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 814
2271 GAAAAACCGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2304
RESULT 12
US-09-338-943-1
Sequence 1, Application US/09338943
Patent No. 637953
GENERAL INFORMATION:
APPLICANT: JOHNSON, KARL F.
APPLICANT: ROTH, STEPHEN
APPLICANT: BUCZALA, STEPHANIE L.
TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO
TITLE OF INVENTION: SACCHARIDE UNITS WITH A POLYGLYCOSYLTRANSFERASE, A
TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE AND GENE ENCODING A
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/338,943
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,140
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Laura A. Coruzzi
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7188-017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8664
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5859 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
ANTI-SENSE: NO
ORIGINAL SOURCE:
STRAIN: F62
FEATURE:
NAME/KEY: CDS
LOCATION: 1..381
FEATURE:
NAME/KEY: CDS
LOCATION: 445..1491
FEATURE:
NAME/KEY: CDS

LOCATION: 2342..3262
FEATURE:
NAME/KEY: CDS
LOCATION: 3322..4335
FEATURE:
NAME/KEY: CDS
LOCATION: 4354..5196
US-09-338-943-1

Query Match 87.7%; Score 726; DB 3; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

1 ATGCAAAACCAAGCTTATGAGCTTACCTTCCGCGCAGAAAGGCGCATTTGCCAT 60
1491 ATGCAAAACCAAGCTTATGAGCTTACCTTCCGCGCAGAAAGGCGCATTTGCCA 1550
61 ACCCTGCGCAGGCAAGCGCATCCCGTTTCACTTTTTCAGCAGCTGATGCCGTGAAAG 120
1551 ACCCTGCGCAGGCAAGCGCATCCCGTTTCACTTTTTCAGCAGCTGATGCCGTGAAAG 1610
121 CTGGAACAGGCAATGCGGCACTGCTCCCGCTTGTCCGCGCAACCCCTATTGAGCGCA 180
1611 CTGGAACAGGCAATGCGGCACTGCTCCCGCTTGTCCGCGCAACCCCTATTGAGCGCA 1670
181 GTGGAAGAAAGCGCTTTTATGAGCAAGCGCTATTGTGGAAGAGGCAATTGACGAAGT 240
1671 GTGGAAGAAAGCGCTTTTATGAGCAAGCGCTATTGTGGAAGAGGCAATTGACGAAGT 1730
241 CTGCGGTATATCAACCGTATTGGAAGCAAGCTTTTATCTCGCGCAAGGTGAGAAATTC 300
1731 GTACCTATATGCGCGTATTGGAAGCAAGCTTTTATCTCGCGCAAGGTGAGAAATTC 1790
301 CTGCGCAAGAGCGCTTGTGCAAGCAAGCTTTGACCCCGGATACCGCTTTATGCTCCG 360
1791 CTGCGCAAGAGCTTGTGCAAGCAAGCTTTGACCCCGGATACCGCTTTATGCTCCG 1850
361 TTGGAAGAGTATTGAGCAAGCTGACCTGCGCCCTCGCGGAGTTCAGTTCAGGAG 420
1851 TTGGAAGAGTATTGAGCAAGCTGACCTGCGCCCTCGCGGAGTTCAGTTCAGGAG 1910
421 CGGCGCTTTTCCGCTGTGGAAGCAAGCACTGCGGAGAGCGCGGCTATATCATTTCCGA 480
1911 CGGCGCTTTTCCGCTGTGGAAGCAAGCACTGCGGAGAGCGCGGCTATATNTTCCGA 1970
481 AAAGGATGCGGTTTCTGCAAGAGTTTCCCGCTCGCGCGCAAGGCGCTGACCCC 540
1971 AAAGGATGCGGTTTCTGCAAGAGTTTCCCGCTCGCGCGCAAGGCGCTGACCCC 2030
541 GTGATCTGATGATGTTGAGCGATTTTTCAGAGGAGAAATGCGGTTTCCAGCTC 600
2031 GTGATCTGATGATGTTGAGCGATTTTTCAGAGGAGAAATGCGGTTTCCAGCTC 2090
601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTCAAGCAAAACAGCGCA 660
2091 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTCAAGCAAAACAGCGCA 2150
661 TTGGGCAAGCTTATGCAAGCAAGCGCTCTGAAACCGCAACAGCAAGGCGCGATTC 720
2151 TTGGGCAAGCTTATGCAAGCAAGCGCTCTGAAACCGCAACAGCAAGGCGCGATTC 2210
721 CCGGCAACAGATTCAAAACAGCGCTGATCCCGCTTGAACCAAAATGAGAGGAGAAAG 780
2211 CCGGCAACAGATTCAAAACAGCGCTGATCCCGCTTGAACCAAAATGAGAGGAGAAAG 2270
781 GAAAAACGCGGCAAGGCGCGCAAGCTTATG 814
2271 GAAAAACGCGGCAAGGCGCGCAAGCTTATG 2304

RESULT 13
US-10-007-267A-1
Sequence 1, Application US/10007267A

Patent No. 6780624
GENERAL INFORMATION:
APPLICANT: Cotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
FILE REFERENCE: 040853-01-5029-02
CURRENT APPLICATION NUMBER: US/10/007,267A
CURRENT FILING DATE: 2001-12-03
PRIORITY APPLICATION NUMBER: US 09/333,412
PRIORITY FILING DATE: 1999-06-15
PRIORITY APPLICATION NUMBER: US 08/878,360
PRIORITY FILING DATE: 1997-06-18
PRIORITY APPLICATION NUMBER: US 08/683,426
PRIORITY FILING DATE: 1996-07-18
PRIORITY APPLICATION NUMBER: US 08/312,387
PRIORITY FILING DATE: 1994-09-24
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 5859
TYPE: DNA
ORGANISM: Neisseria gonorrhoeae
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)-(381)
OTHER INFORMATION: glys (glycyl tRNA synthetase beta chain)
FEATURE:
NAME/KEY: misc.feature
LOCATION: (445)-(1491)
OTHER INFORMATION: 1gca
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2342)-(3262)
OTHER INFORMATION: 1gfc
FEATURE:
NAME/KEY: misc.feature
LOCATION: (3322)-(4335)
OTHER INFORMATION: 1gfd
FEATURE:
NAME/KEY: misc.feature
LOCATION: (4354)-(5196)
OTHER INFORMATION: 1gfe
US-10-007-267A-1
Query Match 87.7%; Score 726; DB 3; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.8e-216;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

1 ATGCAAAACCAAGCTTATGAGCTTACCTTCCGCGCAGAAAGGCGCATTTGCCAT 60
1491 ATGCAAAACCAAGCTTATGAGCTTACCTTCCGCGCAGAAAGGCGCATTTGCCA 1550
61 ACCCTGCGCAGGCAAGCGCATCCCGTTTCACTTTTTCAGCAGCTGATGCCGTGAAAG 120
1551 ACCCTGCGCAGGCAAGCGCATCCCGTTTCACTTTTTCAGCAGCTGATGCCGTGAAAG 1610
121 CTGGAACAGGCAATGCGGCACTGCTCCCGCTTGTCCGCGCAACCCCTATTGAGCGCA 180
1611 CTGGAACAGGCAATGCGGCACTGCTCCCGCTTGTCCGCGCAACCCCTATTGAGCGCA 1670
181 GTGGAAGAAAGCGCTTTTATGAGCAAGCGCTATTGTGGAAGAGGCAATTGACGAAGT 240
1671 GTGGAAGAAAGCGCTTTTATGAGCAAGCGCTATTGTGGAAGAGGCAATTGACGAAGT 1730
241 CTGCGGTATATCAACCGTATTGGAAGCAAGCTTTTATCTCGCGCAAGGTGAGAAATTC 300
1731 GTACCTATATGCGCGTATTGGAAGCAAGCTTTTATCTCGCGCAAGGTGAGAAATTC 1790
301 CTGCGCAAGAGCGCTTGTGCAAGCAAGCTTTGACCCCGGATACCGCTTTATGCTCCG 360
1791 CTGCGCAAGAGCTTGTGCAAGCAAGCTTTGACCCCGGATACCGCTTTATGCTCCG 1850
361 TTGGAAGAGTATTGAGCAAGCTGACCTGCGCCCTCGCGGAGTTCAGTTCAGGAG 420

Db 1851 TTGAAAACGATGTTATGACGTCGACCTCGCCCTCGGCGTGCGGACTACGCGCGG 1910
QY 421 CGGCGCTTTCGCTGTGGAAGCGAACACTGCGGGAACGCGCGGCTATATCATTTCCCGA 480
Db 1911 CGGCGCTTTCGCTGTGGAAGCGAACACTGCGGGAACGCGCGGCTATATCATTTCCCGA 1970
QY 481 AAGCGATGCGGTTTTCCTGACAGTTTCCGCGCTCGCGCGCGAAGGCTGACCCG 540
Db 1971 AAGCGATGCGGTTTTCCTGACAGTTTCCGCGCTCGCGCGCGAAGGCTGACCCG 2030
QY 541 GTCGATCTGATGATGTTTCAGCGATTTTTCGACAGGAAAGAAATGCGGTTTTCAGCTC 600
Db 2031 GTCGATCTGATGATGTTTTCAGCGATTTTTCGACAGGAAAGAAATGCGGTTTTCAGCTC 2090
QY 601 AATCCGCGCTTGCGCGCGCAAGAGCTGATATGCAAGTTTTCAGACCAAAACAGCGCA 660
Db 2091 AATCCGCGCTTGCGCGCGCAAGAGCTGATATGCAAGTTTTCAGACCAAAACAGCGCA 2150
QY 661 TTGGGAGCGCTGATCGAACAGACCGCTCTCTGAAACGCAAAACAGGCGCGATTC 720
Db 2151 TTGGGAGCGCTGATCGAACAGACCGCTCTCTGAAACGCAAAACAGGCGCGATTC 2210
QY 721 CCGGCAACACATTCAAACACCGCGCTGATCCGCGCTTTCAGCAAAATCAGCAGGAAAG 780
Db 2211 CCGGCAACACATTCAAACACCGCGCTGATCCGCGCTTTCAGCAAAATCAGCAGGAAAG 2270
QY 781 GAAAAACGCGCGCAAGGCGCGCAACAGTTCAATTG 814
Db 2271 GAAAAACGCGCGCAAGGCGCGCAACAGTTCAATTG 2304

RESULT 14

US-10-007-267A-7
; Sequence 7, Application US/10007267A
; Patent No. 6780624
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: 040853-01-5029-02
; CURRENT APPLICATION NUMBER: US/10/007,267A
; CURRENT FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent# version 3.2
; SEQ ID NO 7
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: (1491)..(2330)
; OTHER INFORMATION: lgtcb
US-10-007-267A-7

Query Match 87.7%; Score 726; DB 3; Length 5859;

Best Local Similarity 93.2%; Pred. No. 2,8e-216;

Matches /59; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGCAAAACCAAGTATGAGTTAGCTTCGCGCGAAGCAAGGCGGACATTCGCCGAT 60
Db 1491 ATGCAAAACCAAGTATGAGTTAGCTTCGCGCGAAGCAAGGCGGACATTCGCCGAT 1550
QY 61 ACTTCGGAGGACGCGCATCCCGTTTCAGTTTTCAGACGACTGATGCCGCTCTGAAG 120

Db 1551 ACTTCGGAGTTCGCGGATCCCGTTTCAGATTTTTCAGACGACTGATGCCGCTGAAAG 1610
QY 121 CTGGAACAGGCAATGGCGGAACCTGTCGCCGCTTGTCGCGCACCCCTATTTGACGGA 180
Db 1611 CTGGAACAGGCAATGGCGGAACCTGTCGCCGCTTGTCGCGCACCCCTATTTGACGGA 1670
QY 181 GTGAAAAAGCTGCTTTATGACGACCGCGATTTGTGTGGAAGCAGGCAATTGACGAAGT 240
Db 1671 GTGAAAAAGCTGCTTTATGACGACCGCGATTTGTGTGGAAGCAGGCAATTGACGAAGT 1730
QY 241 CTGCGTATATCACCCTGATTTGAGGACGACGTTTACTCGGCGAAGGTGAGGAAAAATTG 300
Db 1731 GTACGTAATATGCGCTGATTTGGAAGATGATGCTTACTCGGGAAGGCGGAGCAGTTTC 1790
QY 301 CTGCGGAAGAGCTGCTGCAAGAAAGCTTTGACCGGATACCGCTTTATGTCGCGC 360
Db 1791 GTACGTAATATGCGCTGATTTGGAAGATGATGCTTACTCGGGAAGGCGGAGCAGTTTC 1850
QY 361 TTGAAAACGATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGCGGATTTACTCGCGG 420
Db 1851 TTGAAAACGATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGCGGACTACGCGCGG 1910
QY 421 CGGCGCTTTCGCTGTGGAAGCGAACACTGCGGGAACGCGCGGCTATATCATTTCCCGA 480
Db 1911 CGGCGCTTTCGCTGTGGAAGCGAACACTGCGGGAACGCGCGGCTATATCATTTCCCGA 1970
QY 481 AAGCGATGCGGTTTTCCTGACAGTTTCCGCGCTCGCGCGCGCAAGGCTGACCCG 540
Db 1971 AAGCGATGCGGTTTTCCTGACAGTTTCCGCGCTCGCGCGCGCAAGGCTGACCCG 2030
QY 541 GTCGATCTGATGATGTTTTCAGCGATTTTTCGACAGGAAAGAAATGCGGTTTTCAGCTC 600
Db 2031 GTCGATCTGATGATGTTTTCAGCGATTTTTCGACAGGAAAGAAATGCGGTTTTCAGCTC 2090
QY 601 AATCCGCGCTTGCGCGCGCAAGAGCTGATATGCAAGTTTTCAGACCAAAACAGCGCA 660
Db 2091 AATCCGCGCTTGCGCGCGCAAGAGCTGATATGCAAGTTTTCAGACCAAAACAGCGCA 2150
QY 661 TTGGGAGCGCTGATCGAACAGACCGCTCTCTGAAACGCAAAACAGGCGCGATTC 720
Db 2151 TTGGGAGCGCTGATCGAACAGACCGCTCTCTGAAACGCAAAACAGGCGCGATTC 2210
QY 721 CCGGCAACACATTCAAACACCGCGCTGATCCGCGCTTTCAGCAAAATCAGCAGGAAAG 780
Db 2211 CCGGCAACACATTCAAACACCGCGCTGATCCGCGCTTTCAGCAAAATCAGCAGGAAAG 2270
QY 781 GAAAAACGCGCGCAAGGCGCGCAACAGTTCAATTG 814
Db 2271 GAAAAACGCGCGCAAGGCGCGCAACAGTTCAATTG 2304

RESULT 15

US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text

```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/09/557,884
3 FILING DATE: 23-Apr-2000
4 CLASSIFICATION: <Unknown>
5
6 PRIOR APPLICATION DATA:
7 APPLICATION NUMBER: 08/476,102
8 FILING DATE: JUN-5-1995
9
10 ATTORNEY/AGENT INFORMATION:
11 NAME: Michelle S. Marks
12 REGISTRATION NUMBER: 41,971
13 REFERENCE/DOCKET NUMBER: PB186P3
14
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 301-309-8504
17 TELEFAX: 301-309-8439
18
19 INFORMATION FOR SEQ ID NO: 1:
20
21 SEQUENCE CHARACTERISTICS:
22 LENGTH: 1830121 base pairs
23 TYPE: nucleic acid
24 STRANDEDNESS: double
25 TOPOLOGY: linear
26
27 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
28
29 US-09-557-884-1

```

Query Match	10.5%;	Score 86.6;	DB 3;	length 1830121;
Best Local Similarity	51.0%;	Pred. No. 3.6e-15;		
Matches 231; Conservative	0;	Mismatches 219;	Indels 3;	Gaps 1;

Qy	180	AGTGAAGAAAAGCGCTGTTATATAGACGACCGCCGATTTGTGTGGAAGCAGGACGATTTGACGCAAGG	239
Db	570747	AGGAGAAAAGCGCTGTTATATATAGCCATTTCTTATATATGAAATATAGTGTGTGATGATGAAAA	5706888
Qy	240	TCCTGCGGTATATACCCGTAATTTTGAGGACGACGTTTTACTCGCGGAAGTGAGGAAAAAT	299
Db	570687	TTTGAATATCTCAAAATTTTTGAGATGATGTAATTTCTGTGGGGAAGATGCGGAATTT	570628
Qy	300	CCTTGCCGGAAGACGCTGGCTGCAAGAACGCTTTGACCCGGATACCGCCTTTATCGTCCG	359
Db	570637	TTTGAACCAAAATGAAATGGTATAAAAACAGGTTTGTATTTAATGATATTTTATATTTGG	570568
Qy	360	CTTGAAGAACGATGTTATATGACAGT---CCTGACCTTGCCCTCGCGGTGCGGATTTACTG	416
Db	570567	TTTAAATACTTTTTTAAACGCCAGTTAAACTTTGAGAAACAAACTTAAATTTCCACCTTTAA	570508
Qy	417	CGGCGCGCGCTTTCCGCTGTGGAAGGCAACATGCGGGGAGCGCGCGCTATATCATTTTC	476
Db	570507	TTCTTAGAACCTTGTATATTTTAAATTAACACTACGTGGGGGAGCGCGGTTATATATTTTC	570448
Qy	477	CGGAAGAGGATGCGCGTTTTCTGTGACAGGTTTTCGCGCCCTGCGCGCCGGAAGGCGCTGCA	536
Db	570447	TCAAGGTGCGGCTAAATATATGTATATGAAATATTTAAAGAAATATTTCTTCTTGATGAAATTTGT	570388
Qy	537	CCCGCGTCAATCTGATGATGATGTCAGCGATTTTTTTCGACAGGGAAGAAATGCGGTTTTGCCA	596
Db	570387	TGCAGTTGATGAACTTAATTTTAAATAAATTAGTGTATGTTGTAATATATATGTTCTATCA	570328
Qy	597	GCTCAATCCCGCTTGTGCGCCCAAGAGCTGCA	629
Db	570327	ACTGATCCAGCAATTTGTATTTCAAGAACTTCA	570295

Search completed: April 7, 2006, 18:22:52
Job time : 305.232 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 18:02:11 ; Search time 1357.05 Seconds
(without alignments)
5045.535 Million cell updates/sec

Title: US-09-211-691-1

Perfect score: 828

Sequence: 1 atgcaaacacacgttaccag.....tcattgtccttccaataa 828

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA Main:
1: /cgn2_6/prodata/1/pubpna/us07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/us08_PUBCOMB.seq:*
3: /cgn2_6/prodata/1/pubpna/us09_PUBCOMB.seq:*
4: /cgn2_6/prodata/1/pubpna/us09_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/us10_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/us10_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/us10_PUBCOMB.seq:*
8: /cgn2_6/prodata/1/pubpna/us10_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/us10_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/us11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	828	100.0	828	US-09-211-691-1	Sequence 1, Appl1
2	828	100.0	828	US-10-317-773-1	Sequence 1, Appl1
3	828	100.0	828	US-10-317-773-1	Sequence 1, Appl1
4	826.4	99.8	46594	US-10-915-740A-4	Sequence 4, Appl1
5	826.4	99.8	2242716	US-10-915-740A-1068	Sequence 1068, Appl1
6	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
7	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
8	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
9	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
10	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
11	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
12	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
13	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
14	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
15	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
16	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
17	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
18	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
19	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
20	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
21	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
22	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1
23	826.4	99.8	5859	US-10-007-267-1	Sequence 1068, Appl1

24	57	6.9	1830121	8	US-10-158-865-1	Sequence 1, Appl1
25	57	6.9	1830121	9	US-10-981-687-1	Sequence 1, Appl1
26	40.2	4.9	1351	7	US-10-767-701-13469	Sequence 13469, A
27	39.8	4.8	4532	10	US-11-097-143-13511	Sequence 13511, A
28	39	4.7	984	10	US-11-097-143-34814	Sequence 34814, A
29	39	4.7	3068	10	US-11-097-143-34813	Sequence 34813, A
30	38.6	4.7	954	6	US-10-369-493-40566	Sequence 40566, A
31	37.8	4.6	2511	7	US-10-451-866-2	Sequence 2, Appl1
32	37.8	4.6	4540	8	US-10-451-866-4	Sequence 4, Appl1
33	37.4	4.5	1229	8	US-10-739-930-5413	Sequence 5413, Ap
34	37.2	4.5	760	5	US-10-027-632-7118	Sequence 7118, Ap
35	37.2	4.5	760	5	US-10-027-632-7119	Sequence 7119, Ap
36	37.2	4.5	760	6	US-10-027-632-7118	Sequence 7118, Ap
37	37.2	4.5	760	6	US-10-027-632-7119	Sequence 7119, Ap
38	36.8	4.4	14643	6	US-10-156-761-931	Sequence 931, App
39	36.8	4.4	31422	7	US-10-204-862A-2	Sequence 2, Appl1
40	36.8	4.4	31422	10	US-11-005-156-2	Sequence 1, Appl1
41	36.8	4.4	9025608	6	US-10-156-761-1	Sequence 2, Appl1
42	36.4	4.4	1122	7	US-10-282-122A-32048	Sequence 32048, A
43	36.4	4.4	7788	6	US-10-329-079-8	Sequence 8, Appl1
44	36.4	4.4	19211	9	US-10-915-740A-54	Sequence 54, Appl1
45	36.4	4.4	37360	6	US-10-329-079-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-09-211-691-1
Sequence 1, Application US/09211691
Patent No. US20020034805A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Fusion Proteins for Use in Enzymatic Synthesis of
FILE REFERENCE: 019957-012910US
CURRENT APPLICATION NUMBER: US/09/211,691
CURRENT FILING DATE: 1998-12-14
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 828
TYPE: DNA
ORGANISM: Neisseria meningitidis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(828)
OTHER INFORMATION: beta-1,4-galactosyltransferase (1gfb)
US-09-211-691-1

Query Match 100.0%; Score 828; DB 3; Length 828;
Best Local Similarity 100.0%; Pred. No. 6.6e-26;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCAAAACAGTATACGTTACCTCCGCGAAGACGAGGCGACATTCGCGAT	60
DB	1	ATGCAAAACAGTATACGTTACCTCCGCGAAGACGAGGCGACATTCGCGAT	60
QY	61	ACCTGGGAGGACGAGCAGCAGTCCGTTTTCAGATTTTCAGCAGTATGCGTGAAGG	120
DB	61	ACCTGGGAGGACGAGCAGCAGTCCGTTTTCAGATTTTCAGCAGTATGCGTGAAGG	120
QY	121	CTGGAACGAGGACGAGCAGCAGTCCGTTTTCAGATTTTCAGCAGTATGCGTGAAGG	180
DB	121	CTGGAACGAGGACGAGCAGCAGTCCGTTTTCAGATTTTCAGCAGTATGCGTGAAGG	180
QY	181	GTGGAACGAGGACGAGCAGCAGTCCGTTTTCAGATTTTCAGCAGTATGCGTGAAGG	240
DB	181	GTGGAACGAGGACGAGCAGCAGTCCGTTTTCAGATTTTCAGCAGTATGCGTGAAGG	240

```
Db 181 GTGAAAAAGCCCTGTTATGAGCAGCCGTAATTGTGAAAGAGGCAATTGAGCAAGGT 240
QY 241 CTGCGGTATATCAACCGTATTTTGAAGACAGAGCTTTTAACTCGGCGAAGGTGAGAAAAATTC 300
Db 241 CTGCGGTATATCAACCGTATTTTGAAGACAGAGCTTTTAACTCGGCGAAGGTGAGAAAAATTC 300
QY 301 CTGCGAAGACGCTTGAGTGAAGAACGCTTGACCCCGGTAACCGGCTTTATGTCGCG 360
Db 301 CTGCGAAGACGCTTGAGTGAAGAACGCTTGACCCCGGTAACCGGCTTTATGTCGCG 360
QY 361 TTGGAACCATGTTTATGACAGTCTGACCTGCGCTCGGCGTGGCGATTAATGCGG 420
Db 361 TTGGAACCATGTTTATGACAGTCTGACCTGCGCTCGGCGTGGCGATTAATGCGG 420
QY 421 CGGCGCTTTCGCTGTGGAAGAGCAACAGTGGGAGAGCGGCGGCTATATCATTTGCCGA 480
Db 421 CGGCGCTTTCGCTGTGGAAGAGCAACAGTGGGAGAGCGGCGGCTATATCATTTGCCGA 480
QY 481 AAAGCGATGCGGTTTTCCTGGAAGAGTTCGCGCTGCGCGCGAAGGCTGACCCG 540
Db 481 AAAGCGATGCGGTTTTCCTGGAAGAGTTCGCGCTGCGCGCGAAGGCTGACCCG 540
QY 541 GTGATCTGATGATGTTTGAAGAGCTGATTTTTCGACAGGGAAGAAATGCGGTTGCGAGTC 600
Db 541 GTGATCTGATGATGTTTGAAGAGCTGATTTTTCGACAGGGAAGAAATGCGGTTGCGAGTC 600
QY 601 AATCCCGGCTTGCGCGCCCAAGAGCTGATTAATGCAAGTTTTCAGACCAAAACAGCGCA 660
Db 601 AATCCCGGCTTGCGCGCCCAAGAGCTGATTAATGCAAGTTTTCAGACCAAAACAGCGCA 660
QY 661 TTGGGAGAGCTGATTCGAACAGACCGCTCTGGAACCGGAACAGAAAGCGGATTC 720
Db 661 TTGGGAGAGCTGATTCGAACAGACCGCTCTGGAACCGGAACAGAAAGCGGATTC 720
QY 721 CCGGCAACACATTCGAAACCGCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAG 780
Db 721 CCGGCAACACATTCGAAACCGCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAG 780
QY 781 GAAAAAGCGCGGCAAGGCGCAACAGTTCATTTGCTTTCCAAATTA 828
Db 781 GAAAAAGCGCGGCAAGGCGCAACAGTTCATTTGCTTTCCAAATTA 828

RESULT 2
US-10-317-773-1
; Sequence 1, Application US/10317773
; Publication No. US20030180928A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Wakachuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Fusion Protein Comprising a UDP-GalNAc 4' Epimerase and a
; FILE REFERENCE: 019633-000812US
; CURRENT APPLICATION NUMBER: US/10/317,773
; CURRENT FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 09/211,691
; PRIOR FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 1
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(828)
; OTHER INFORMATION: beta-1,4-galactosyltransferase (1gcb)
US-10-317-773-1
```

```
Query Match 100.0%; Score 828; DB 6; Length 828;
Best Local Similarity 100.0%; Pred. No. 6,66-267;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAAAACCAAGCTTATGAGCTTACCTCCGCGGAGAAACGAGGCGCCACATTCGCAT 60
Db 1 ATGCAAAACCAAGCTTATGAGCTTACCTCCGCGGAGAAACGAGGCGCCACATTCGCAT 60
QY 61 ACCCTGAGGAGCAACGCGATCCGTTTCAAGTTTTCAGAGCACTGATGCGCTGAAAAG 120
Db 61 ACCCTGAGGAGCAACGCGATCCGTTTCAAGTTTTCAGAGCACTGATGCGCTGAAAAG 120
QY 121 CTGGAACAGGCAATGAGGAGAACTGTCGCCGCTGTGCGGCAACCCCTATTTGAGCGGA 180
Db 121 CTGGAACAGGCAATGAGGAGAACTGTCGCCGCTGTGCGGCAACCCCTATTTGAGCGGA 180
QY 181 GTGAAAAAGCCCTGCTTATGAGCAAGCCGTAATTTGGAAGACAGGCAATTGACAGAGT 240
Db 181 GTGAAAAAGCCCTGCTTATGAGCAAGCCGTAATTTGGAAGACAGGCAATTGACAGAGT 240
QY 241 CTGCGGTATATCAACCGTATTTTGAAGACAGCTTTTAACTCGGCGAAGGTGAGAAAAATTC 300
Db 241 CTGCGGTATATCAACCGTATTTTGAAGACAGCTTTTAACTCGGCGAAGGTGAGAAAAATTC 300
QY 301 CTGCGCAAGACGCTTGAGTGAAGACGCTTTGACCCGGAATCCGCTTTATGCTCGCG 360
Db 301 CTGCGCAAGACGCTTGAGTGAAGACGCTTTGACCCGGAATCCGCTTTATGCTCGCG 360
QY 361 TTGGAACCATGTTTATGACAGTCTGACCTGCGCTCGGCGAGGAGGATTAATGCGG 420
Db 361 TTGGAACCATGTTTATGACAGTCTGACCTGCGCTCGGCGAGGAGGATTAATGCGG 420
QY 421 CGGCGCTTTCGCTGTGGAAGAGCAACAGTGGGAGAGCGGCTATATCATTTCCGGA 480
Db 421 CGGCGCTTTCGCTGTGGAAGAGCAACAGTGGGAGAGCGGCTATATCATTTCCGGA 480
QY 481 AAAGCGATGCGGTTTTCCTGGAAGAGTTCGCGCTGCGCGCGAAGGCTGACCCG 540
Db 481 AAAGCGATGCGGTTTTCCTGGAAGAGTTCGCGCTGCGCGCGAAGGCTGACCCG 540
QY 541 GTGATCTGATGATGTTTGAAGAGCTGATTTTTCGACAGGGAAGAAATGCGGTTGCGAGTC 600
Db 541 GTGATCTGATGATGTTTGAAGAGCTGATTTTTCGACAGGGAAGAAATGCGGTTGCGAGTC 600
QY 601 AATCCCGGCTTGCGCGCCCAAGAGCTGATTAATGCAAGTTTTCAGACCAAAACAGCGCA 660
Db 601 AATCCCGGCTTGCGCGCCCAAGAGCTGATTAATGCAAGTTTTCAGACCAAAACAGCGCA 660
QY 661 TTGGGAGAGCTGATTCGAACAGACCGCTCTGGAACCGGAACAGAAAGCGGATTC 720
Db 661 TTGGGAGAGCTGATTCGAACAGACCGCTCTGGAACCGGAACAGAAAGCGGATTC 720
QY 721 CCGGCAACACATTCGAAACCGCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAG 780
Db 721 CCGGCAACACATTCGAAACCGCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAG 780
QY 781 GAAAAAGCGCGGCAAGGCGCAACAGTTCATTTGCTTTCCAAATTA 828
Db 781 GAAAAAGCGCGGCAAGGCGCAACAGTTCATTTGCTTTCCAAATTA 828

RESULT 3
US-10-317-428-1
; Sequence 1, Application US/10317428
; Publication No. US20030186414A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Wakachuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Nucleic Acid That Encodes a Fusion Protein
; FILE REFERENCE: 019633-00081US
; CURRENT APPLICATION NUMBER: US/10/317,428
```

```

; CURRENT FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 09/211,691
; PRIOR FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(828)
; OTHER INFORMATION: beta-1,4-galactosyltransferase (lgtB)
US-10-317-428-1

Query Match      100.0%; Score 828; DB 6; Length 828;
Best Local Similarity 100.0%; Pred. No. 6.6e-267;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGCAAAACCAAGTATACGTTAGCTTCCGCGGAGAACGACGGGCGACATTGCCGAT 60
DB      1 ATGCAAAACCAAGTATACGTTAGCTTCCGCGGAGAACGACGGGCGACATTGCCGAT 60

QY      61 ACCTTGGGAGGAGCGGACATCCGGTTTCACTTTTGGAGCGACATGACCGCTTGAAAG 120
DB      61 ACCTTGGGAGGAGCGGACATCCGGTTTCACTTTTGGAGCGACATGACCGCTTGAAAG 120

QY      121 CTGGAACAGGCAATGCGGAACTGTCGCCGGCTTGCGGCGACCCCTATTGAGCGGA 180
DB      121 CTGGAACAGGCAATGCGGAACTGTCGCCGGCTTGCGGCGACCCCTATTGAGCGGA 180

QY      122 CTGGAACAGGCAATGCGGAACTGTCGCCGGCTTGCGGCGACCCCTATTGAGCGGA 180
DB      122 CTGGAACAGGCAATGCGGAACTGTCGCCGGCTTGCGGCGACCCCTATTGAGCGGA 180

QY      181 GTGGAAGAAAGCTGCTTTATGACCCACGCCGATTTGGAAGAGGATTTGAAGAGGT 240
DB      181 GTGGAAGAAAGCTGCTTTATGACCCACGCCGATTTGGAAGAGGATTTGAAGAGGT 240

QY      241 CTGCGCGTATATCAACGATTTTGAAGAGAGTTTACTCGGCGAAGTGAAGAAATTC 300
DB      241 CTGCGCGTATATCAACGATTTTGAAGAGAGTTTACTCGGCGAAGTGAAGAAATTC 300

QY      301 CTGCGCGAAGACGCTTGCGTGAAGAACGCTTTGACCCCGGATPACCGCTTTATCGTCCG 360
DB      301 CTGCGCGAAGACGCTTGCGTGAAGAACGCTTTGACCCCGGATPACCGCTTTATCGTCCG 360

QY      361 TTGGAACAGATGTTTATGCAAGTCTGACCTGCGCTCGGCGTGGCGGATTACTGCGGG 420
DB      361 TTGGAACAGATGTTTATGCAAGTCTGACCTGCGCTCGGCGTGGCGGATTACTGCGGG 420

QY      421 CCGGCTTTCCGCTGTTGGAAGAACGACACTGGGGGACGCGGGCTATATCATTTCCGA 480
DB      421 CCGGCTTTCCGCTGTTGGAAGAACGACACTGGGGGACGCGGGCTATATCATTTCCGA 480

QY      481 AAAGCATCGGCTTTTCTGACAGAGTTTGCGCGCTGCGCGCGGAGAGGCTGACCCC 540
DB      481 AAAGCATCGGCTTTTCTGACAGAGTTTGCGCGCTGCGCGCGGAGAGGCTGACCCC 540

QY      541 GTGCGATCTATGATGTTTCAAGCGATTTTTCGACAGGGAAGAAATGCGGTTTGCGAGCTC 600
DB      541 GTGCGATCTATGATGTTTCAAGCGATTTTTCGACAGGGAAGAAATGCGGTTTGCGAGCTC 600

QY      601 AATCCCGCTTGCGCGCGCGGACGCTGACATATGCAAGTTTTCAGACCAAAACAGCGCA 660
DB      601 AATCCCGCTTGCGCGCGCGGACGCTGACATATGCAAGTTTTCAGACCAAAACAGCGCA 660

QY      661 TTGGGAGAGCTGATCGAACAGACGCGCTCTCTGAACCGCAACAGGAGGCGCATTC 720
DB      661 TTGGGAGAGCTGATCGAACAGACGCGCTCTCTGAACCGCAACAGGAGGCGCATTC 720

QY      721 CCGCGCAACATTTAAACACGCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG 780
DB      721 CCGCGCAACATTTAAACACGCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG 780
```

```

QY      781 GAAAGCGCGGCAAGGCGGCAAGACGTTCAATTGCTTCCAAATTA 828
DB      781 GAAAGCGCGGCAAGGCGGCAAGACGTTCAATTGCTTCCAAATTA 828

RESULT 4
US-10-915-740A-4/C
; Sequence 4, Application US/10915740A
; Publication No. US2005019136A1
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tettelein, Herre
; APPLICANT: Venter, J. Craig
; APPLICANT: Massignani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroza
; APPLICANT: Ratti, Giulio
; APPLICANT: Scariello, Maria
; APPLICANT: Scariello, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizzi, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: 002441.00090
; CURRENT APPLICATION NUMBER: US/10/915,740A
; PRIOR FILING DATE: 2004-08-11
; PRIOR APPLICATION NUMBER: 09/806,866
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: USSN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USSN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 46594
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-915-740A-4

Query Match      99.8%; Score 826.4; DB 9; Length 46594;
Best Local Similarity 99.9%; Pred. No. 1.1e-265;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ATGCAAAACCAAGTATACGTTAGCTTCCGCGGAGAACGACGGGCGACATTGCCGAT 60
DB      33537 ATGCAAAACCAAGTATACGTTAGCTTCCGCGGAGAACGACGGGCGACATTGCCGAT 33478

QY      61 ACCTTGGGAGGAGCGGACATCCGGTTTCACTTTTGGAGCGACATGACCGCTTGAAAG 120
DB      33477 ACCTTGGGAGGAGCGGACATCCGGTTTCACTTTTGGAGCGACATGACCGCTTGAAAG 33418

QY      121 CTGGAACAGGCAATGCGGAACTGTCGCCGGCTTGCGGCGACCCCTATTGAGCGGA 180
DB      33417 CTGGAACAGGCAATGCGGAACTGTCGCCGGCTTGCGGCGACCCCTATTGAGCGGA 33358

QY      181 GTGGAAGAAAGCTGCTTTATGACAGCGCGGATTTGTGGAAGACAGGCAATTGACGAAGT 240
DB      33357 GTGGAAGAAAGCTGCTTTATGACAGCGCGGATTTGTGGAAGACAGGCAATTGACGAAGT 33298

QY      241 CTGCGGTATATCAACGATTTTGAAGAGACGTTTACTCGGGAAGGTGAGGAAATTC 300
DB      33297 CTGCGGTATATCAACGATTTTGAAGAGACGTTTACTCGGGAAGGTGAGGAAATTC 33238

QY      301 CTGCGCGAAGACGCTTGCGTGAAGACGCTTTGAACCGGATPACCGCTTTATCGTCCG 360
DB      33237 CTGCGCGAAGACGCTTGCGTGAAGACGCTTTGAACCGGATPACCGCTTTATCGTCCG 33178

QY      361 TTGGAACAGATGTTTATGCAAGTCTGACCTGCGCTCGGCGGCGGATTACTGCGGG 420
```

```
DB 33117 TTGGAACGATGTTATGACGTCGACCTCGCCCTCGGCGGATTCGCGG 33118
QY 421 CGGCGCTTCCGCTGTTGAAAGCAACACTGGGGGACGCGGGCTATATCATTTCCGA 480
DB 33117 CGGCGCTTCCGCTGTTGAAAGCAACACTGGGGGACGCGGGCTATATCATTTCCGA 33058
QY 481 AAAGCAGTGGGTTTTCCTGCAAGATTGCGCCCTGCGCGCAAGGGCTGCAAGG 540
DB 33057 AAAGCAGTGGGTTTTCCTGCAAGATTGCGCCCTGCGCGCAAGGGCTGCAAGG 32998
QY 541 GTGATCTGATGATGTTGACGATTTTTCGACAGGAAAGAAATGCGGTTTCCGAGCTC 600
DB 32997 GTGATCTGATGATGTTGACGATTTTTCGACAGGAAAGAAATGCGGTTTCCGAGCTC 32938
QY 601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTCATTCAGCAAGTTCCAGCAAGAGCA 660
DB 32937 AATCCGCGCTTGTGCGCCCAAGAGCTGATTCATTCAGCAAGTTCCAGCAAGAGCA 32878
QY 661 TTGGGAGAGCTGATTCGAAACGACGCGCTTCTGTAACCGCAAGCAAGGCGGATTC 720
DB 32877 TTGGGAGAGCTGATTCGAAACGACGCGCTTCTGTAACCGCAAGCAAGGCGGATTC 32818
QY 721 CCCGCAACACATTCAAACACGCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAG 780
DB 32817 CCCGCAACACATTCAAACACGCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAG 32758
QY 781 GAAAAAGCGCGCAAGAGCGCGCAAGCTTCATTTGCTTCATTA 828
DB 32757 GAAAAAGCGCGCAAGAGCGCGCAAGCTTCATTTGCTTCATTA 32710
```

```
RESULT 5
US-10-915-740A-1068/C
/ Sequence 1068, Application US/10915740A
/ Publication No. US20050191316A1
/ GENERAL INFORMATION:
/ APPLICANT: Frazer, Claire M.
/ APPLICANT: Hickey, Brian
/ APPLICANT: Peterson, Jeremy
/ APPLICANT: Tetteelin, Hervé
/ APPLICANT: Venter, J. Craig
/ APPLICANT: Masignani, Vega
/ APPLICANT: Galeotti, Cesira
/ APPLICANT: Ratti, Giulio
/ APPLICANT: Scarselli, Maria
/ APPLICANT: Scariato, Vincenzo
/ APPLICANT: Rappuoli, Rino
/ APPLICANT: Pizzi, Mariagrazia
/ APPLICANT: Grandi, Guido
/ TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
/ FILE REFERENCE: 002441.00090
/ CURRENT APPLICATION NUMBER: US/10/915,740A
/ PRIOR FILING DATE: 2004-08-11
/ PRIOR APPLICATION NUMBER: 09/806,866
/ PRIOR FILING DATE: 1999-10-08
/ PRIOR APPLICATION NUMBER: USSN 60/103,794
/ PRIOR FILING DATE: 1998-10-09
/ PRIOR APPLICATION NUMBER: USSN 60/132,068
/ PRIOR FILING DATE: 1999-04-30
/ PRIOR APPLICATION NUMBER: PCT/US99/25373
/ PRIOR FILING DATE: 1999-10-08
/ NUMBER OF SEQ ID NOS: 1068
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1068
/ LENGTH: 2242716
/ TYPE: DNA
/ ORGANISM: Neisseria meningitidis
US-10-915-740A-1068
```

Query Match 99.8%; Score 826.4; DB 9; Length 2242716;
Best Local Similarity 99.9%; Pred. No. 4.6e-265;

```
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCAAAACCAAGTTATGAGCTTACCTCCGCGGAAACGAGGCGCAATTCGCGAT 60
DB 1997047 ATGCAAAACCAAGTTATGAGCTTACCTCCGCGGAAACGAGGCGCAATTCGCGAT 1996988
QY 61 ACCCTCGGACGACAGGATCCGTTGATTTTTCAGAGCACTGATGCTCGTGAAGG 120
DB 1996987 ACCCTCGGACGACAGGATCCGTTGATTTTTCAGAGCACTGATGCTCGTGAAGG 1996928
QY 121 CTGGAACAGGCAATGCGGAACTGTCGCCGCTTGTGCGGACCCCTATTTGAGCGGA 180
DB 1996927 CTGGAACAGGCAATGCGGAACTGTCGCCGCTTGTGCGGACCCCTATTTGAGCGGA 1996868
QY 181 GTGAAAAAGCGCTGTTATGAGCAAGCGCGGATTTGTGAAGCAGGATTTGAAGAGT 240
DB 1996867 GTGAAAAAGCGCTGTTATGAGCAAGCGCGGATTTGTGAAGCAGGATTTGAAGAGT 1996808
QY 241 CTGCGGTATATCACCGTATTTGAGAGCAGCTTTTACTGCGCGAAGGTGAGGAAATTC 300
DB 1996807 CTGCGGTATATCACCGTATTTGAGAGCAGCTTTTACTGCGCGAAGGTGAGGAAATTC 1996748
QY 301 CTGCGCAAGAGCGCTTGTGAGAGAGAGCTTTGACCGCGATACCGCTTATGTCGCG 360
DB 1996747 CTGCGCAAGAGCGCTTGTGAGAGAGAGCTTTGACCGCGATACCGCTTATGTCGCG 1996688
QY 361 TTGGAACGATGTTTATGACAGTCTGACCTTGCCCTTCGCGGCGTGGCGGATTTACTGCGGG 420
DB 1996687 TTGGAACGATGTTTATGACAGTCTGACCTTGCCCTTCGCGGCGTGGCGGATTTACTGCGGG 1996628
QY 421 CGGCGCTTCCGCTGTTGAAAGGAAACACTGGGGGACGCGGGCTATATCATTTCCGA 480
DB 1996627 CGGCGCTTCCGCTGTTGAAAGGAAACACTGGGGGACGCGGGCTATATCATTTCCGA 1996568
QY 481 AAAGGATGCGGATTTTCTGAGACAGGTTTGCCTGCGCGCGCGCAAGGCGCTCACCCC 540
DB 1996567 AAAGGATGCGGATTTTCTGAGACAGGTTTGCCTGCGCGCGCGCAAGGCGCTCACCCC 1996508
QY 541 GTGATCTGATGATGTTTCAGAGATTTTTCGACAGGAAAGAAATGCGGTTTCCAGCTC 600
DB 1996507 GTGATCTGATGATGTTTCAGAGATTTTTCGACAGGAAAGAAATGCGGTTTCCAGCTC 1996448
QY 601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTCATTCAGCAAGTTCCAGCAAGAGCGCA 660
DB 1996447 AATCCGCGCTTGTGCGCCCAAGAGCTGATTCATTCAGCAAGTTCCAGCAAGAGCGCA 1996388
QY 661 TTGGGACAGCTGATTCGAAACAGACCGGCTCTGTAACGCAAAACAGCAAGCGCGATTC 720
DB 1996387 TTGGGACAGCTGATTCGAAACAGACCGGCTCTGTAACGCAAAACAGCAAGCGCGATTC 1996328
QY 721 CCCGCAACACATTCAAACACGCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAGG 780
DB 1996327 CCCGCAACACATTCAAACACGCGCTGATCCGCGCTTGAACCAAAATCAGAGGAAAGG 1996268
QY 781 GAAAAAGCGCGCAAGAGCGCGCAAGCTTCATTTGCTTCATTA 828
DB 1996267 GAAAAAGCGCGCAAGAGCGCGCAAGCTTCATTTGCTTCATTA 1996220
```

```
RESULT 6
US-10-007-267-1
/ GENERAL INFORMATION:
/ APPLICANT: Gottschlich, Emil C.
/ TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
/ OLIGOSACCHARIDES, AND GENES ENCODING THEM
/ NUMBER OF SEQUENCES: 12
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Klauber & Jackson
/ STREET: 411 Hackensack Avenue
/ CITY: Hackensack
/ STATE: New Jersey
/ COUNTRY: USA
/ ZIP: 07601
```


COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,267
FILING DATE: 03-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/333,412
FILING DATE: 15-Jun-1999
APPLICATION NUMBER: 08/312,387
FILING DATE: July 7, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-007-267-1

Query Match 87.7%; Score 726; DB 5; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.5e-232;

Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

1 ATGCAAAACCAAGTATACGTTAGCTTCCGCGCAAGACGAGGCGCATTCGCAT 60
1491 ATGCAAAACCAAGTATACGTTAGCTTCCGCGCAAGACGAGGCGCATTCGC 1550
61 ACCCTGGCGAGGACGCGCATCCGCTTCACTTTTGGACGCACTGATCCGCTGA 120
1551 ACCCTGGCGAGTGGCGCATCCGCTTCACTTTTGGACGCACTGATCCGCTGA 1610
121 CTGGAACAGGCAATGGCGCAACTCGTCCGCGCTGCGCGGCAACCCCTATTGA 180
1611 CTGGAACAGGCAATGGCGCAACTCGTCCGCGCTGCGCGGCAACCCCTATTGA 1670
181 GTGAAAAAAGCTGCTTATGACGCAAGCGCTATTTGGAAGAGGATGGAAGAGT 240
1671 GTGAAAAAAGCTGCTTATGACGCAAGCGCTATTTGGAAGAGGATGGAAGAGT 1730
241 CTGCGCTATATACCGTATTGAGAGACGCTTTTACTCGGCGAAGGTGAAGAAATTC 300
1731 GTACCGTATATCCCGCTATTTGGAAGATGATGCTTACTCGGCGAAGGCGGAGCAGTTC 1790
301 CTGCGCGAAGACGCTTGGCTGCAAGAACGCTTGAACCGGATACCGCTTATTCGTC 360
1791 CTGCGCGAAGATCTTGGCTGCAAGAACGCTTGAACCGGATACCGCTTATTCGTC 1850
361 TTGAAAAAGATGTTATGACGCTGCAAGCTCGCCCTCGCGCGATGAGATTACTG 420
1851 TTGAAAAAGATGTTATGACGCTGCAAGCTCGCCCTCGCGCGATGAGATTACTG 1910
421 CGGCGCTTTCGCTGTTGGAAGACGACCTGCGGAGCGCGGCTATATCATTTCCGA 480
1911 CGGCGCTTTCGCTGTTGGAAGACGACCTGCGGAGCGCGGCTATATCATTTCCGA 1970
481 AAGCGATGCGGTTTTCCTGCAAGGTTTGGCGCCCTGCGCGCAAGAGGCTGACCCC 540
1971 AAGCGATGCGGTTTTCCTGCAAGGTTTGGCGCCCTGCGCGCAAGAGGCTGACCCC 2030
541 GTGATCTGATGATGTTTTCGCAAGGTTTTCGCAAGGTTTTCGCAAGGTTTTCGCA 600
2031 GTGATCTGATGATGTTTTCGCAAGGTTTTCGCAAGGTTTTCGCAAGGTTTTCGCA 2090
601 AATCCCGCTTGGCGCGCAAGAGCTGCAATTAAGCAAGTTTTCGCAAGCAAGCGCA 660
2091 AATCCCGCTTGGCGCGCAAGAGCTGCAATTAAGCAAGTTTTCGCAAGCAAGCGCA 2150

661 TTGCGAGCGCTATGCAACGACGCGCTCTGTAACCGCAAAACGAGGCGCATTC 720
2151 TTGCGAGCGCTATGCAACGACGCGCTCTGTAACCGCAAAACGAGGCGCATTC 2210
721 CCGCGCAACATTCATCAACGCGCTGATCCGCGCTTGAACCAATTCGAGGAGAA 780
2211 CCGCGCAACATTCATCAACGCGCTGATCCGCGCTTGAACCAATTCGAGGAGAA 2270
781 GAAAAACGCGCGCAAGGCGCGCAACGTTGATTC 814
2271 GAAAAACGCGCGCAAGGCGCGCAACGTTGATTC 2304

RESULT 7

US-10-007-267-7

GENERAL INFORMATION:

APPLICANT: Gottschlich, Emil C.

TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSER: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/007,267

FILING DATE: 03-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/333,412

FILING DATE: 15-Jun-1999

APPLICATION NUMBER: 08/312,387

FILING DATE: July 7, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-095

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-007-267-7

Query Match 87.7%; Score 726; DB 5; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2.5e-232;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

1 ATGCAAAACCAAGTATACGTTAGCTTCCGCGCAAGACGAGGCGCATTCGCCAT 60
1491 ATGCAAAACCAAGTATACGTTAGCTTCCGCGCAAGACGAGGCGCATTCGCCAT 1550
61 ACCCTGGCGAGGACGCGCATCCGCTTCACTTTTGGACGCACTGATCCGCTGA 120
1551 ACCCTGGCGAGTGGCGCATCCGCTTCACTTTTGGACGCACTGATCCGCTGA 1610
121 CTGGAACAGGCAATGGCGCAACTCGTCCGCGCTGCGCGCAACCCCTATTGA 180
1611 CTGGAACAGGCAATGGCGCAACTCGTCCGCGCTGCGCGCAACCCCTATTGA 1670
181 GTGAAAAAAGCTGCTTATGACGCAAGCGCTATTTGGAAGAGGATGGAAGAGT 240
1671 GTGAAAAAAGCTGCTTATGACGCAAGCGCTATTTGGAAGAGGATGGAAGAGT 1730
241 CTGCGCTATATACCGTATTGAGAGACGCTTTTACTCGGCGAAGGTGAAGAAATTC 300

```
Db 1731 G|A|C|C|T|A|T|A|T|G|C|C|G|T|A|T|T|G|A|A|G|A|T|G|T|T|A|C|T|G|G|C|A|A|A|G|C|G|G|A|G|C|A|G|T|T|C 1790
Qy 301 C|T|T|G|C|G|A|A|G|A|C|G|T|T|G|C|T|G|C|A|A|G|A|C|G|T|T|T|A|C|C|G|G|A|T|A|C|G|C|T|T|T|A|T|G|T|C|C|G 360
Db 1791 C|T|T|G|C|G|A|A|G|A|T|C|T|T|G|C|T|G|C|A|A|G|A|C|G|T|T|T|G|A|C|C|C|G|A|T|T|C|C|G|C|T|T|T|G|T|G|T|C|C|G 1850
Qy 361 T|T|G|A|A|A|C|A|T|G|T|T|A|T|G|A|C|G|T|C|G|A|C|G|T|C|C|C|C|T|C|G|G|C|G|T|G|G|G|A|T|T|A|T|C|T|G|C|G|G 420
Db 1851 T|T|G|A|A|A|C|A|T|G|T|T|A|T|G|A|C|G|T|C|G|A|C|G|T|C|C|C|C|T|C|G|G|C|G|T|G|G|G|A|T|T|A|T|C|T|G|C|G|G 1910
Qy 421 C|G|G|C|C|T|T|T|C|C|G|T|T|G|A|A|G|C|G|A|A|C|T|G|G|G|G|A|C|G|C|G|G|C|T|A|T|A|T|T|C|C|G|A 480
Db 1911 C|G|G|C|C|T|T|T|C|C|G|T|T|G|A|A|G|C|G|A|A|C|T|G|G|G|G|A|C|G|C|G|G|C|T|A|T|A|T|T|T|C|C|G|A 1970
Qy 481 A|A|G|G|C|A|T|G|G|T|T|T|T|C|T|G|A|C|A|G|T|T|T|C|C|G|C|C|T|C|C|C|C|G|G|A|A|G|G|C|T|G|C|A|C|C|C 540
Db 1971 A|A|G|G|C|A|T|G|G|T|T|T|T|C|T|G|A|C|A|G|T|T|T|C|C|G|C|C|T|C|C|C|C|G|G|A|A|G|G|C|T|G|C|A|C|C|C 2030
Qy 541 G|T|C|G|A|T|C|T|A|T|G|T|T|T|C|A|G|C|G|A|T|T|T|T|T|C|A|C|A|G|G|A|A|G|A|T|G|C|G|T|T|G|C|A|G|T|C 600
Db 2031 G|T|C|G|A|T|T|A|T|A|T|G|T|T|C|G|C|A|A|C|C|T|G|A|C|A|G|G|A|A|G|A|T|G|C|G|T|T|G|C|A|G|T|C 2090
Qy 601 A|A|T|C|C|C|C|C|T|T|G|G|C|C|C|C|C|A|A|G|C|T|G|C|A|T|T|A|T|G|C|A|A|G|T|T|T|C|A|C|A|C|A|A|A|C|A|G|C|G|A 660
Db 2091 A|A|T|C|C|C|C|C|T|T|G|G|C|C|C|C|C|A|A|G|C|T|G|C|A|T|T|A|T|G|C|A|A|G|T|T|T|C|A|C|A|C|A|A|A|C|A|G|C|G|A 2150
Qy 661 T|T|G|G|G|A|G|C|C|T|A|T|C|A|A|C|A|C|G|C|C|T|C|T|G|A|A|C|C|G|A|A|C|A|G|A|A|G|G|C|G|A|T|T|C 720
Db 2151 T|T|G|G|G|A|G|C|C|T|A|T|C|A|A|C|A|C|G|C|C|C|G|C|T|G|A|A|C|C|G|A|A|C|A|G|A|A|T|G|G|C|G|A|T|T|C 2210
Qy 721 C|C|C|G|C|A|A|C|A|T|T|C|A|A|C|A|C|G|C|C|T|G|A|C|C|A|A|A|T|C|A|G|C|A|G|G|A|A|G 780
Db 2211 C|C|G|C|C|A|A|C|A|T|T|C|A|A|C|A|C|G|C|C|T|G|A|C|C|G|C|C|T|T|G|A|C|A|A|A|T|C|G|C|A|G|G|A|A|G 2270
Qy 781 G|A|A|A|A|C|G|C|G|C|A|A|G|C|G|C|G|C|A|A|C|A|G|T|T|C|A|T|T|G 814
Db 2271 G|A|A|A|A|C|G|C|G|C|A|A|G|C|G|C|G|C|A|A|C|A|G|T|T|A|T|C|G 2304
```

```
RESULT 8
US-10-096-129-1
; Sequence 1, Application US/10096129
; Publication No. US20030207406A1
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, KARL F.
; APPLICANT: ROTH, STEPHEN
; APPLICANT: BUCZALA, STEPHANIE L.
; TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO SACCHARIDE UNITS WITH A
; FILE REFERENCE: POLYGLYCOSYLTRANSFERASE
; CURRENT APPLICATION NUMBER: US/10/096,129
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/338,943
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: US 08/478,140
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-096-129-1
```

```
Query Match 87.7%; Score 726; DB 6; Length 5859;
Best Local Similarity 93.2%; Pred. No. 2,5e-232;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
Qy 1 ATGCAGAAACCAAGTATACGCTTAGCTTCCGCGGAGAAAGAGGCGGACATTTGCCAT 60
Db 1491 ATGCAGAAACCAAGTATACGCTTAGCTTCCGCGGAGAAAGAGGCGGACATTTGCCA 1550
```

```
Qy 61 A|C|T|T|G|C|G|A|G|C|A|C|G|G|A|T|C|C|G|T|T|T|C|A|G|C|A|T|G|A|T|C|C|G|T|C|T|G|A|A|A|G 120
Db 1551 A|C|T|T|G|C|G|A|G|C|A|C|G|G|A|T|C|C|G|T|T|T|C|A|G|C|A|T|G|A|T|C|C|G|T|C|T|G|A|A|A|G 1610
Qy 121 C|T|G|A|A|C|A|G|G|C|A|A|T|G|G|C|G|A|A|C|T|G|T|C|C|G|G|C|T|T|G|G|G|C|A|C|C|C|T|A|T|T|G|A|G|G|A 180
Db 1611 C|T|G|A|A|C|A|G|G|C|A|A|T|G|G|C|G|A|A|C|T|G|T|C|C|G|G|C|T|T|G|G|G|C|A|C|C|C|T|A|T|T|G|A|G|G|A 1670
Qy 181 G|T|G|A|A|A|A|G|C|T|G|T|T|A|T|A|G|C|A|C|G|C|G|T|A|T|T|G|A|A|G|C|A|G|C|A|T|T|G|A|C|A|A|G|T 240
Db 1671 G|T|G|A|A|A|A|G|C|T|G|T|T|A|T|A|G|C|A|C|G|C|G|T|A|T|T|G|A|A|G|C|A|G|C|A|T|T|G|A|C|A|A|G|C 1730
Qy 241 C|T|G|C|G|T|A|T|C|A|C|C|G|T|T|T|G|A|G|A|C|A|C|G|T|T|T|A|C|T|G|G|G|A|A|G|G|A|A|A|A|T|T|C 300
Db 1731 G|T|A|C|G|T|A|T|C|C|G|T|A|T|T|T|G|A|A|G|A|T|A|T|T|C|T|A|C|T|G|G|G|A|A|G|C|G|G|A|G|C|A|G|T|T|C 1790
Qy 301 C|T|T|G|C|G|A|A|G|A|C|G|T|T|G|C|T|G|C|A|A|A|C|G|T|T|T|G|A|C|C|G|G|A|T|A|C|G|C|T|T|A|T|G|T|C|C|G|C 360
Db 1791 C|T|T|G|C|G|A|A|G|A|T|A|C|T|T|G|C|T|G|C|A|A|A|C|G|T|T|T|G|A|C|C|C|G|A|T|T|C|G|T|C|G|T|C|C|G|C 1850
Qy 361 T|T|G|A|A|A|G|A|T|A|T|A|T|A|C|A|G|T|C|T|G|A|C|T|G|C|C|C|C|C|G|G|G|G|G|G|G|A|T|T|A|C|T|G|C|G|G 420
Db 1851 T|T|G|A|A|A|G|A|T|A|T|A|T|A|C|A|G|T|C|T|G|A|C|T|G|C|C|C|C|C|G|G|G|G|G|G|G|A|T|T|A|C|T|G|C|G|G 1910
Qy 421 C|G|G|C|C|T|T|T|C|G|T|T|G|A|A|G|C|A|C|A|C|T|G|G|G|G|A|C|G|G|G|G|C|T|A|T|A|T|C|A|T|T|T|C|C|G|A 480
Db 1911 C|G|G|C|C|T|T|T|C|G|T|T|G|A|A|G|C|A|C|A|C|T|G|G|G|G|A|C|G|G|G|G|C|T|A|T|A|T|T|T|C|C|G|A 1970
Qy 481 A|A|G|G|A|T|G|G|T|T|T|T|C|T|G|A|C|A|G|T|T|T|C|C|G|C|C|C|C|G|G|G|G|G|G|G|A|A|G|G|C|T|G|C|A|C|C|C 540
Db 1971 A|A|G|G|A|T|G|G|T|T|T|T|C|T|G|A|C|A|G|T|T|T|C|C|G|C|C|C|C|G|G|G|G|G|G|G|A|A|G|G|C|T|G|C|A|C|C|C 2030
Qy 541 G|T|C|G|A|T|C|T|A|T|G|T|T|T|C|A|G|C|G|A|T|T|T|T|T|C|A|C|A|G|G|A|A|G|A|T|G|C|G|T|T|G|C|A|G|T|C 600
Db 2031 G|T|C|G|A|T|T|A|T|A|T|G|T|T|C|G|C|A|A|C|C|T|G|A|C|A|C|A|G|G|A|A|G|A|T|G|C|G|T|T|G|C|A|G|T|C 2090
Qy 601 A|A|T|C|C|C|C|C|T|T|G|G|C|C|C|C|C|A|A|G|C|T|G|C|A|T|T|A|T|G|C|A|A|G|T|T|T|C|A|C|A|A|A|C|A|G|C|G|A 660
Db 2091 A|A|T|C|C|C|C|C|T|T|G|G|C|C|C|C|C|A|A|G|C|T|G|C|A|T|T|A|T|G|C|A|A|G|T|T|T|C|A|C|A|A|A|C|A|G|C|G|A 2150
Qy 661 T|T|G|G|G|A|G|C|C|T|A|T|C|A|A|C|A|C|G|C|C|T|C|T|G|A|A|C|C|G|A|A|C|A|G|A|A|G|G|C|G|A|T|T|C 720
Db 2151 T|T|G|G|G|A|G|C|C|T|A|T|C|A|A|C|A|C|G|C|C|C|G|C|T|G|A|A|C|C|G|A|A|C|A|G|A|A|T|G|G|C|G|A|T|T|C 2210
Qy 721 C|C|C|G|C|A|A|C|A|T|T|C|A|A|C|A|C|G|C|C|T|G|A|C|C|G|C|C|T|T|G|A|C|A|A|A|T|C|A|G|C|A|G|G|A|A|G 780
Db 2211 C|C|G|C|C|A|A|C|A|T|T|C|A|A|C|A|C|G|C|C|T|G|A|C|C|G|C|C|T|T|G|A|C|A|A|A|T|C|G|C|A|G|G|A|A|G 2270
Qy 781 G|A|A|A|A|C|G|C|G|C|A|A|G|C|G|C|G|C|A|A|C|A|G|T|T|C|A|T|T|G 814
Db 2271 G|A|A|A|A|C|G|C|G|C|A|A|G|C|G|C|G|C|A|A|C|A|G|T|T|A|T|C|G 2304
```

```
RESULT 9
US-10-654-528-1
; Sequence 1, Application US/10654528
; Publication No. US20040043464A1
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
```

Qy	1	ATGGAACACGCGTATTCAGCTAGCTTCGCGCGAAGAACGAGGGCGGCACTATGGCGAT	60
Db	1491	ATGCAAAACCAAGTATTCAGCTTGGCTTCGCGCGAAGAACGAGGGCGGCACTATGGCGGA	1550
Qy	61	ACCTTGGGCAAGGACGCGCATCCCGTTACGATTTTTCAGACGCACTGATGCGCTCTGAAG	120
Db	1551	ACCTTGGGCAAGTGCAGCGCATCCCGTTCAAGTTTTCAGACGCACTGATGCGCTCTGAAG	1610
Qy	121	CTGGAACAGGCATATGCGCGAATCTGCTCCCGGCTTGTGGCGGACCCCTATTGACCGGA	180
Db	1611	CTGGAACAGGCATATGCGCGAATCTGCTCCCGGCTTGTGGCGGACCCCTATTGACCGGA	1670
Qy	181	GTGGAAAAAGCTGCTTTATGAGCCACGCGATTTGTGGAAGCAGGCACTTGACGAAAGT	240
Db	1671	GTGGAAAAAGCTGCTTTATGAGCCACGCGATTTGTGGAACAGGCACTTGACGAAAGG	1730
Qy	241	CTGCGCGTATATCAACCGTATTGAGACGACGTTTACTCGCGAAGGTGAGAAAAATTC	300
Db	1731	GTACCGTATATCCCGTATTGAAAGTATGCTTACTCGGCGAAGGCGGAGCGAGTTC	1790

QY 301 CTGGCGAAGACGCTTGGCTGCAAGACGCTTTGACCCGAGTACCGCTTTTATGTCGCC 360
DB 1791 CTGGCGAAGATCTTGGCTGCAAGACGCTTTGACCCGAGTACCGCTTTTATGTCGCC 1850
QY 361 TTGGAAACGATGTTTATGACGCTCTGACCTCGCCCTCGGCGGTGGGAGATTACTGCGGG 420
DB 1851 TTGGAAACGATGTTTATGACGCTCTGACCTCGCCCTCGGCGGTGGGAGATTACTGCGGG 1910
QY 421 CGGCGCTTTCGCGTGTGGAAAGCGAACATCTGGGCGACGCGCGCTTATATCATTTCCCA 480
DB 1911 CGGCGCTTTCGCGTGTGGAAAGCGAACATCTGGGCGACGCGCGCTTATATCATTTCCCA 1970
QY 481 AAAAGCATGCGGTTTTCCTGCAAGAGTTTGGCGCCCTGCGCCGCAAGGCGTGCACCC 540
DB 1971 AAGGCGATGCGGTTTTCCTGCAAGAGTTTGGCGCCCTGCGCCGCAAGGCGTGCACCC 2030
QY 541 GTGATCTGATGATGTTTTCAGCGGATTTTTCGACAGGAAAGAAATGCGGTTTTCAGCTC 600
DB 2031 GTGATCTGATGATGTTTTCAGCGGATTTTTCGACAGGAAAGAAATGCGGTTTTCAGCTC 2090
QY 601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTTATGCCAAGTTTCACGACCAAAACAGCGCA 660
DB 2091 AATCCGCGCTTGTGCGCCCAAGAGCTGATTTATGCCAAGTTTCACGACCAAAACAGCGCA 2150
QY 661 TTGGGAGCGCTGATGCAACGACGCGCTCTGAAACGCGAAACGAAAGGCGCGATTC 720
DB 2151 TTGGGAGCGCTGATGCAACGACGCGCTCTGAAACGCGAAACGAAAGGCGCGATTC 2210
QY 721 CCGCGCAACATTTCAAACCGCGCTGATTCGCGCTTGAACCAAAATCAGCAGGAAAG 780
DB 2211 CCGCGCAACATTTCAAACCGCGCTGATTCGCGCTTGAACCAAAATCAGCAGGAAAG 2270
QY 781 GAAAAACGCGCGCAAGGCGCGCAAGCTTCAATTG 814
DB 2271 GAAAAACGCGCGCAAGGCGCGCAAGCTTCAATTG 2304

RESULT 11
US-10-472-260-175
; Sequence 175, Application US/10472260
; Publication No. US20040265328A1
; GENERAL INFORMATION:
; APPLICANT: MICROBIOLOGICAL RESEARCH AUTHORITY
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS LIMITED
; FILE REFERENCE: GMS/DJC/23480
; CURRENT APPLICATION NUMBER: US/10/472,260
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: Patent version 3.1
; SEQ ID NO 175
; LENGTH: 840
; TYPE: DNA
; ORGANISM: *Neisseria meningitidis* (group B)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(840)
; OTHER INFORMATION: NMB1926
US-10-472-260-175

Query Match 51.4%; Score 426; DB 8; Length 840;
Best Local Similarity 75.2%; Pred. No. 6, 6e-132;
Matches 531; Conservative 0; Mismatches 175; Indels 0; Gaps 0;
QY 1 ATGCAAAACCAAGTATACGTTAGCTTCGCGGAGAACGAGGCGCAGATTGCCCAT 60
DB 1 ATGCAAAACCAAGTATACGTTAGCTTCGCGGAGAACGAGGCGCAGATTGCCCAT 60
QY 61 ACCCTGGAGGACGCGCATCCCGTTGCTTTTTCAGAGCATGATGCGCTGAAAG 120
DB 61 ACCCTGGAGGACGCGCATCCCGTTGCTTTTTCAGAGCATGATGCGCTGAAAG 120

QY 121 CTGGAACAGGCAATGGCGGAATCTGTCGCCGCTTGTGCGGCAACCCCTATTGACGCGA 180
DB 121 CTGGAATCGGATGATGGCGGAATTTGTGCGCGGATTTGGCAAAACAGCACCTGTGAGCGGA 180
QY 181 GTGGAAAAAGCTGCTTTATAGAGCAAGCCGATTTGTTGAAAGCAGGCAATGGAAGT 240
DB 181 GTGGAAAAAGCTGCTTTATAGAGCAAGCCGATTTGTTGAAAGCAGGCAATGGAAGT 240
QY 241 CTGCGTATATACCGTATTTGAGGACAGCTTTTACTCGCGGAAAGTGAAGAAAAATTC 300
DB 241 CTGCGCTATGTTGCGTATTTGAGGATGATGCTGCTGGCAAGAGCGAAGAAAGTTTC 300
QY 301 CTGGCGAAGACGCTTGGCTGCAAGACGCTTTGACCGGATACCGCTTATGTCGCG 360
DB 301 CTGGCGAAGATCTTGGCTTGAAGAGGCTTTGATTAAGATTCGCTTTTATGCTTGT 360
QY 361 TTGGAAACGATGTTTATGACGCTCTGACCTCGCCCTCGGCGTGGCGGATTTACTGCGGG 420
DB 361 TTGGAAACGATGTTTATGCAAGTTATGTGACCGGATTAAGTCTGAAATTAAGAGAT 420
QY 421 CGGCGCTTTCGCTGTTGAAAGCGAACACTGGGGGAGCGGCGCTATATCATTTCCCGA 480
DB 421 CGGCTATTTCTTGTGCTGAGAGCGAACATTTGGGGAGCGCTGCTATATCATTTCCGCT 480
QY 481 AAGCGATGCGGTTTTCCTGCAAGTTTGGCGCCCTGCGCCGCAAGGCGTGCACCC 540
DB 481 AAGCGATGCGGTTTTCCTGCAAGTTTGGCGCCCTGCGCCGCAAGGCGTGCACCC 540
QY 541 GTGATCTGATGATGTTTTCAGCGGATTTTTCGACAGGAAAGAAATGCGGTTTTCAGCTC 600
DB 541 GTGATCTGATGATGTTTATCTTATTTCTTGTATAGAGGAGGAGTCTGTTTATCAGGTT 600
QY 601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTTATGCCAAGTTTCACGACCAAAACAGCGCA 660
DB 601 AATCCGCGCTTATGACCAAGAAATGATTTATGCAAGTTTCACGATTAAGAAACAGATG 660
QY 661 TTGGGAGCGCTGATGCAACGACGCGCTCTGAAACGCGCAACAGC 706
DB 661 TTGGGATGAGTATTTGAAAAAGATAGGGAACAGAAAGAAAGACACC 706

RESULT 12
US-10-795-159-570
; Sequence 570, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: US/10/795,159
; PRIOR FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: Patent version 3.2
; SEQ ID NO 570
; LENGTH: 14547
; TYPE: DNA
; ORGANISM: *H. influenzae*
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14474)..(14474)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14508)..(14508)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14518)..(14518)
; OTHER INFORMATION: n = a, c, g, or t
US-10-795-159-570

Query Match 11.2%; Score 93; DB 9; Length 14547;
Best Local Similarity 51.9%; Pred. No. 1.3e-19;
Matches 225; Conservative 0; Mismatches 215; Indels 3; Gaps 1;

QY 180 AGGAGAAAAAGCTGCTTTATGAGCAAGCCGATTTGTGAGACAGGCAATTGACAGAGG 239
DB 11201 AGGAGAAAAAGGCTGTTTATTTATGAGCAATTTCTATTATGAAATGATGTAAGAAAA 11260

QY 240 TCGGCGGTATATACCGGTATTTAGAGACAGGCTTTACTCGGAGAGGTGAGAAAAAT 299
DB 11261 TTTGGAATATCTACAAATTTTGAAGATGATGTAATTTCTGCGAGAAATGCGAAGTGT 11320

QY 300 CCTTGGCCAGACGCTTGTGCGAGAACGCTTTGACCCGATATACCGCTTTATCGTCG 359
DB 11321 TTTGGACACAGATGAAATGTTAAAAACGTTTGAATTTTAAATGATTTTATTTATGCG 11380

QY 360 CTTGGAAAGCATTTTATGACGT---CTGACCTCGCCCTCGGCGGTGGCGGATTAATG 416
DB 11381 TTTAGAACTTTTTCAGCCAGTTAACTTGAGAAAACAAACTAAATTCACCTTTTAA 11440

QY 417 CGGCGCGCGCTTCCGCTGTGAGAACGAAACGTCGGGGAGCGCGGCTATATCATTTTC 476
DB 11441 TTTCTAGAACTTTGATATTTTAAATTCAGCTACGCGGGAGCGCTGTTATATTTATTC 11500

QY 477 CCGAAAGCGATCGGCTTTTCTGGAACAGTTTGCCTCGCGCCGCGGAGGCGTGA 536
DB 11501 TCGAGGTGCGGCTTAATATGTAATTTGAATTTTAAAGAAATTTCTTGTGATGAATTTGT 11560

QY 537 CCCCCTGATCGATGATGATGATTCAGCAATTTTTCAGACAGGAAGAAATGCGGTTGCCA 596
DB 11561 TCGAGTGTATGAATCTATTATTTTAAATTAAGTTGATGATGTAATTAATTTATTTATCA 11620

QY 597 GCTCAATCCCGCTTGTGCGCCCAAGAGCTGCA 629
DB 11621 ACTTAATCCAGCAATTTGATTCAGAACTCCA 11653

RESULT 13
US-10-795-159-681
; Sequence 681, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 681
; LENGTH: 106645
; TYPE: DNA
; ORGANISM: H. influenzae
US-10-795-159-681

Query Match 11.2%; Score 93; DB 9; Length 106645;
Best Local Similarity 51.9%; Pred. No. 2.7e-19;
Matches 225; Conservative 0; Mismatches 215; Indels 3; Gaps 1;

QY 180 AGGAGAAAAAGCTGCTTTATGAGCAAGCCGATTTGTGAGACAGGCAATTGACAGAGG 239
DB 98869 AGGAGAAAAAGGCTGTTTATTTATGAGCAATTTCTATTATGAAATGATGTAAGAAAA 98928

QY 240 TCGGCGGTATATACCGGTATTTAGAGACAGGCTTTACTCGGAGAGGTGAGAAAAAT 299
DB 98929 TTTGGAATATCTACAAATTTTGAAGATGATGTAATTTCTTGGGAGAAATGCGAAGTGT 98988

QY 300 CCTTGGCCAGACGCTTGTGCGAGAACGCTTTGACCCGATATACCGCTTTATCGTCG 359
DB 98989 TTTGGACACAGATGAAATGTTAAAAACGTTTGAATTTTAAATGATATTTTATTTATGCG 99048

QY 360 CTTGGAAGCATGTTTATGACGT---CTGACCTCGCCCTCGGCGGTGGCGGATTAATG 416
DB 99049 TTTAGAACTTTTATTTTACGCCAGTTAACTTGAGAAACAAATTTCACTTTTAA 99108

QY 417 CGGCGCGCGCTTTCGCTGTTGAAACGAAACATCGGGGAGCGCGGCTATATCATTTTC 476
DB 99109 TTTAGAACTTTGATATTTTAAATTCAGCTACGCGGGAGCGGCTGTTATTTATTTTC 99168

QY 477 CCGAAAGCATGCGGCTTTTCTGGAACAGTTTCCGCGCCCTCGGCGCGGAGGCTGA 536
DB 99169 TCAAGGTGCGGCTAAATATGTAATTTTAAAGAAATATTTCTTCTGATGAATTTGT 99228

QY 537 CCCCCTGATCGATGATGATGATTCAGCAATTTTTCAGACAGGAAGAAATGCGGTTGCA 596
DB 99229 TCGAGTGTATGAATCTATTATTTTAAATTAATTAATTAATTAATTAATTTATGATG 99288

QY 597 GCTCAATCCCGCTTGTGCGCCCAAGAGCTGCA 629
DB 99289 ACTTAATCCAGCAATTTGATTCAGAACTCCA 99321

RESULT 14
US-10-329-670-1/c
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fra
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; PRIOR FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:

```
NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10091)..(10091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
```

```

; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152530) .. (152530)
;
Query Match      10.5%; Score 86.6; DB 7; Length 1830121;
Best Local Similarity 51.0%; Pred. No. 1,1e-16;
Matches 231; Conservative 0; Mismatches 219; Indels 3; Gaps 1;
;
QY 180 AGTGAAGAAAAGCTGCTTTATGAGCCAGCCGCTATGTTGAGACAGGCAATTGACGAGG 239
    |||||
DB 570747 AGGAGAAAAGGCTGTTATTTAGCCATTTCTTATTAATGAAATGATGATGAAA 570688
    |||||
QY 240 TCTGCCGTAATATGACCGTATTTGAGACGATTTTACTGCGGAGAGTGAAGAAAATT 239
    |||||
DB 570687 TTTGTAATATCTCAAAATTTTGAAGATGATTAATCTTGCGAGAAATCGAGATTT 570628
    |||||
QY 300 CCTTGCAGAAAGCCTTGCTGCAAGAACGCTTTGACCCCGATPACCGCTTTATCGTCG 359
    |||||
DB 570627 TTTGAACCAAAATGAATGTTAAACAGTTTGAATTAATGATATTTTATTAATGCG 570568
    |||||
QY 360 CTTGGAAGACATGTTATGACGT---CTGACCTCGCCCTCGGCGTGGCGGATTACTG 416
    |||||
DB 570567 TTTAGAAACTTTTTCAGCCAGTTAACTTGAGAAAAGAACTTAAATTCACCTTTTAA 570508
    |||||
QY 417 CGGCGCGGCTTCCGCTGTTGGAAGCAACACTGGGGGAGCGGCGGCTATATCAATTTC 476
    |||||
DB 570507 TTTCTAGAACTTTGATATTTTAAATCACTGAGGGAGCGAGTTATATTTTC 570448
    |||||
QY 477 CCGAAAAGCGATCGGTTTTTCTGACAGGTTTGCGCGCCCGCCGAGAGGCTGCA 536
    |||||
DB 570447 TCAAGGTGGGCTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 570388
    |||||
QY 537 CCCGTCGATCTGATGATGTTCCAGCATTTTTCGACAGGAAAGAAATGCGGTTGCCA 556
    |||||
DB 570387 TGCAGTGTGATGAACTTATTTTATTAATTAATTAATTAATTAATTAATTAATTA 570328
    |||||
QY 597 GCTCATCCCGCTTGCGGCGCAAGAGCTGCA 629
    |||||
DB 570327 ACTGAATCCAGCAATTTGATTTCAAGACTCCA 570295
    |||||

RESULT 15
US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Publication No. US20040203093A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747) .. (4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921) .. (9921)
; OTHER INFORMATION: n equals a,t,c, or g

;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150) .. (10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298) .. (29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543) .. (36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551) .. (36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636) .. (36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40810) .. (40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416) .. (44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905) .. (44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975) .. (44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732) .. (45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036) .. (47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334) .. (51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602) .. (51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786) .. (51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805) .. (51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369) .. (55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309) .. (65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
```

NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10091)..(10091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature

LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)

Query Match 10.5%; Score 86.6; DB 8; Length 1830121;
Best Local Similarity 51.0%; Pred. No. 1.1e-16;
Matches 231; Conservative 0; Mismatches 219; Indels 3; Gaps 1;

QY 180 AGTGGAAAAAGCGCTGTTATGAGCCGATATGGAAGCAGGATGGAAGG 239
DB 570747 AGGAGAAAAAGCGTGTATTAATGACCTTATTAATGAGATGATGATGAAAA 570688
QY 240 TCTGCCGTATATCAACCGATTTGAGAGACGTTTACTCGCGGAAGTGAAGAAAT 299
DB 570687 TTTGAAATATCTCAAAATTTTGAAGATGATATCTTGGCGGAATTCGAAAGTAT 570628
QY 300 CCTTGCAGAAAGCGCTTGGCTGCAAGAAAGCTTTGACCCGATACCGCTTTATGCTCG 359
DB 570627 TTTGAACCAAAATGATGATGATGATGATGATGATGATGATGATGATGATG 570568
QY 360 CTTGAAAGCAATGTTATGACGCT--CCTGACCTGCGCCGCGGCGATTAATCAG 416
DB 570567 TTTGAACCTTTTACAGCCGATTAATGAGAAACAACTAAATTCACCTTTTAA 570508
QY 417 CGGGCGCGCTTTCGCTGTTGGAAGCACTGGGAGACGGCGGCTATATCATTTTC 476
DB 570507 TTTCTAGGAACCTTGTATTTTAAATCACTCAGGGGAGCGGAGTTATATTTTC 570448
QY 477 CCGAAAGGATGCGGTTTTCCTGAGACAGTTTGCCTGCGCCCGGAGAGGCTGCA 536
DB 570447 TCAAGTCCGCTAATATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 570388
QY 537 CCCCGTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 596
DB 570387 TGCAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 570328
QY 597 GCTCAATCCCGCTTGTGCGCCCAAGAGCTGCA 629
DB 570327 ACTGATCCAGCAATTTGATTAATCAAGACTCA 570295

Search completed: April 7, 2006, 18:56:47
Job time : 1367.05 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr 11 7, 2006, 18:05:06 ; Search time 966.604 Seconds
(without alignments)
3426.906 Million cell updates/sec

Title: US-09-211-691-1

Perfect score: 828
Sequence: 1 atgcaaacacacgtatcag.....tcattgtgccttcacataa 828

Scoring table: IDENTITY NTC
Gapop 10.0 , Gapext 1.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA New:*
1: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /SIDS5/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq:*
7: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq:*
8: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq:*
9: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq:*
10: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq:*
11: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq:*
12: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq:*
13: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq:*
14: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq:*
15: /SIDS5/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	726	87.7	5859	14 US-11-102-497-1	Sequence 1, Appli
2	726	87.7	5859	14 US-11-102-497-7	Sequence 7, Appli
3	721.2	87.1	837	8 US-10-467-657-1309	Sequence 1309, Ap
4	471.8	57.0	840	8 US-10-467-657-1325	Sequence 1325, Ap
5	57	6.9	2449	14 US-11-194-246-26	Sequence 26, Appl
6	38	4.6	1760	14 US-11-024-959-37	Sequence 37, Appl
7	34.6	4.2	453	8 US-10-997-437A-9	Sequence 9, Appli
8	34.6	4.2	1218	8 US-10-750-185-56685	Sequence 56685, A
9	34.6	4.2	1218	8 US-10-750-623-56685	Sequence 56685, A
10	33.8	4.1	11864	9 US-10-330-773-826	Sequence 826, App
11	33.6	4.1	535	9 US-09-925-065A-52784	Sequence 52784, A
12	33.6	4.1	535	9 US-10-301-480-154022	Sequence 154022,
13	33.6	4.1	535	10 US-10-301-480-767431	Sequence 767431,
14	33.6	4.1	7364	14 US-11-096-281-10	Sequence 10, Appl
15	33.4	4.0	626	6 US-09-925-065A-525973	Sequence 525973,
16	33.4	4.0	626	6 US-09-925-065A-525974	Sequence 525974,
17	33.4	4.0	626	6 US-09-925-065A-717818	Sequence 717818,
18	33	4.0	411	8 US-10-467-657-6075	Sequence 6075, Ap

ALIGNMENTS

```
C 19 33 4.0 555 8 US-10-467-657-6055 Sequence 6055, Ap
C 20 33 4.0 555 8 US-10-467-657-7535 Sequence 7535, Ap
C 21 33 4.0 938 11 US-11-096-568A-18718 Sequence 18718, A
C 22 32.8 4.0 1866 8 US-10-467-657-1345 Sequence 1345, A
C 23 32.8 3.9 1694969 7 US-10-506-454-1690 Sequence 1690, Ap
C 24 32.4 3.9 1230 8 US-10-467-657-4857 Sequence 4857, Ap
C 25 32.2 3.9 6021 14 US-11-136-527-274 Sequence 274, App
C 26 32.2 3.9 1878 8 US-10-467-657-6425 Sequence 6425, Ap
C 27 32.2 3.9 1878 8 US-10-467-657-7617 Sequence 7617, Ap
C 28 32 3.9 834 8 US-10-750-185-37550 Sequence 37550, A
C 29 32 3.9 834 8 US-10-750-623-37550 Sequence 37550, A
C 30 32 3.9 1890 8 US-10-750-185-61161 Sequence 61161, A
C 31 32 3.9 1890 8 US-10-750-623-61161 Sequence 61161, A
C 32 32 3.9 2205 8 US-10-467-657-37299 Sequence 3729, Ap
C 33 32 3.9 2205 8 US-10-467-657-6989 Sequence 6989, Ap
C 34 31.8 3.8 808 11 US-11-096-568A-12049 Sequence 12049, A
C 35 31.8 3.8 1263 8 US-10-858-730-149 Sequence 149, App
C 36 31.6 3.8 1999 11 US-11-096-568A-14677 Sequence 14677, A
C 37 31.6 3.8 4734 14 US-11-136-527-2518 Sequence 2518, A
C 38 31.4 3.8 861 14 US-11-137-465-15 Sequence 15, Appl
C 39 31.4 3.8 165857 14 US-11-121-086-34 Sequence 34, Appl
C 40 31.2 3.8 1026 8 US-10-750-185-57052 Sequence 57052, A
C 41 31.2 3.8 1026 8 US-10-750-623-57052 Sequence 57052, A
C 42 31.2 3.8 7011 14 US-11-136-527-3740 Sequence 3740, Ap
C 43 31.2 3.8 9695 14 US-11-096-281-12 Sequence 12, Appl
C 44 31.2 3.8 88421 14 US-11-205-109-1 Sequence 1, Appl
C 45 31 3.7 1348 9 US-10-838-616-33 Sequence 33, Appl

RESULT 1
US-11-102-497-1 ; Sequence 1, Application US/1102497
; Publication No. US20050271690A1
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C
; TITLE OF INVENTION: GENES ENCODING THEM
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/11/102,497
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(381)
; OTHER INFORMATION: glys (glycyl tRNA synthetase beta chain)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (445)..(1491)
; OTHER INFORMATION: 19CA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2342)..(3262)
; OTHER INFORMATION: 19rc
```

```
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (3322)..(4335)
/ OTHER INFORMATION: 1gtd
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (4354)..(5196)
/ OTHER INFORMATION: 1gce
US-11-102-497-1
```

```
Query Match      87.7%; Score 726; DB 14; Length 5859;
Best Local Similarity 93.2%; Pred. No. 3e-215;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
```

```
QY 1 ATGCAAAACCAAGTATACGCTTAGCTTCGCGCGAGAAACGCGGCGGACATTCGCGAT 60
DB 1491 ATGCAAAACCAAGTATACGCTTAGCTTCGCGCGAGAAACGCGGCGGACATTCGCGCA 1550
QY 61 ACCTTCGCGAGGCAAGGCAATCCGCTTTCAGTTTTCAGAGCACTGATGCCGCTGAAAG 120
DB 1551 ACCTTCGCGAGGCAAGGCAATCCGCTTTCAGTTTTCAGAGCACTGATGCCGCTGAAAG 1610
QY 121 CTGGAAACAGGCAATGCGGCAATCGTCCCGGCTTCGCGGCAACCCCTATTTCAGCGGA 180
DB 1611 CTGGAAACAGGCAATGCGGCAATCGTCCCGGCTTCGCGGCAACCCCTATTTCAGCGGA 1670
QY 181 GTGGAAAAAGCGCTTATAGAGCGCAAGCGCTATTTGTGAAGAGGCAATTCGAAAGT 240
DB 1671 GTGGAAAAAGCGCTTATAGAGCGCAAGCGCTATTTGTGAAGAGGCAATTCGAAAGT 1730
QY 241 CTGCGGTATATACCGGTATTTGAGAGCAAGTTTATCTCGCGAAGGTGAGAAAAATTC 300
DB 1731 GTACCGGTATATACCGGTATTTGAGAGCAAGTTTATCTCGCGAAGGTGAGAAAAATTC 1790
QY 301 CTTCGCGAAGACGCTTGGCTGCAAGACGCTTTCAGCCGGAATACCGCTTTATTCGCGGC 360
DB 1791 CTTCGCGAAGACGCTTGGCTGCAAGACGCTTTCAGCCGGAATACCGCTTTATTCGCGGC 1850
QY 361 TTGGAAACGATGTTTATAGACGCTTCGACCTCGCCCTCGGCGGTGCGGATTAATGCGGG 420
DB 1851 TTGGAAACGATGTTTATAGACGCTTCGACCTCGCCCTCGGCGGTGCGGATTAATGCGGG 1910
QY 421 CGGCGCTTTCGCTGTTTGAAGAGCAACTGCGGCGGACGCGGCGCTATATATCTCCGA 480
DB 1911 CGGCGCTTTCGCTGTTTGAAGAGCAACTGCGGCGGACGCGGCGCTATATATCTCCGA 1970
QY 481 AAAGCGATGCGGTTTCTTGAAGAGGTTTGCGCCCTCGCCCGGACGCGCTGCAACCC 540
DB 1971 AAAGCGATGCGGTTTCTTGAAGAGGTTTGCGCCCTCGCCCGGACGCGCTGCAACCC 2030
QY 541 GTTCGATCTGATGATGTTTTCAGCGCATTTTTCAGACAGGAAAGAAATGCGGTTTTCGAGCTC 600
DB 2031 GTTCGATCTGATGATGTTTTCAGCGCATTTTTCAGACAGGAAAGAAATGCGGTTTTCGAGCTC 2090
QY 601 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTTCAGACCAACCAAAACGCGGA 660
DB 2091 AATCCGCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTTCAGACCAACCAAAACGCGGA 2150
QY 661 TTGGGAGGCTGATGATGCAACGAGCGCTCTGGAACCGCAACAGCAAGGCGGCAATTC 720
DB 2151 TTGGGAGGCTGATGATGCAACGAGCGCTCTGGAACCGCAACAGCAAGGCGGCAATTC 2210
QY 721 CCGCGCAACATTCGCAACAGCGCGCTGATCCGCGCTTTCAGCAAAATCAGCAGGAAAGG 780
DB 2211 CCGCGCAACATTCGCAACAGCGCGCTGATCCGCGCTTTCAGCAAAATCAGCAGGAAAGG 2270
QY 781 GAAAAACGCGGCAAGGCGGCAAGGCTTCAATG 814
DB 2271 GAAAAACGCGGCAAGGCGGCAAGGCTTCAATG 2304
```

RESULT 2
US-11-102-497-7

```
/ Sequence 7, Application US/11102497
/ Publication No. US20050271690A1
/ GENERAL INFORMATION:
/ APPLICANT: Gotschlich, Emil C.
/ TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
/ FILE REFERENCE: 040853-01-5029-03/1102,497
/ CURRENT FILING DATE: 2005-04-08
/ PRIOR APPLICATION NUMBER: US/10/654,528
/ PRIOR FILING DATE: 2003-09-02
/ PRIOR APPLICATION NUMBER: US 10/007,267
/ PRIOR FILING DATE: 2001-12-03
/ PRIOR APPLICATION NUMBER: US 09/333,412
/ PRIOR FILING DATE: 1999-06-15
/ PRIOR APPLICATION NUMBER: US 08/878,360
/ PRIOR FILING DATE: 1997-06-18
/ PRIOR APPLICATION NUMBER: US 08/683,426
/ PRIOR FILING DATE: 1996-07-18
/ PRIOR APPLICATION NUMBER: US 08/312,387
/ PRIOR FILING DATE: 1994-09-24
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 7
/ LENGTH: 5859
/ TYPE: DNA
/ ORGANISM: Neisseria gonorrhoeae
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1491)..(2330)
/ OTHER INFORMATION: 1gcb
US-11-102-497-7
```

```
Query Match      87.7%; Score 726; DB 14; Length 5859;
Best Local Similarity 93.2%; Pred. No. 3e-215;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
```

```
QY 1 ATGCAAAACCAAGTATACGCTTAGCTTCGCGCGAGAAACGCGGCGGACATTCGCGAT 60
DB 1491 ATGCAAAACCAAGTATACGCTTAGCTTCGCGCGAGAAACGCGGCGGACATTCGCGCA 1550
QY 61 ACCTTCGCGAGGCAAGGCAATCCGCTTTCAGTTTTCAGAGCACTGATGCCGCTGAAAG 120
DB 1551 ACCTTCGCGAGGCAAGGCAATCCGCTTTCAGTTTTCAGAGCACTGATGCCGCTGAAAG 1610
QY 121 CTGGAAACAGGCAATGCGGCAATCGTCCCGGCTTCGCGGCAACCCCTATTTCAGCGGA 180
DB 1611 CTGGAAACAGGCAATGCGGCAATCGTCCCGGCTTCGCGGCAACCCCTATTTCAGCGGA 1670
QY 181 GTGGAAAAAGCGCTTATAGAGCGCAAGCGCTATTTGTGAAGAGGCAATTCGAAAGT 240
DB 1671 GTGGAAAAAGCGCTTATAGAGCGCAAGCGCTATTTGTGAAGAGGCAATTCGAAAGT 1730
QY 241 CTGCGGTATATACCGGTATTTGAGAGCAAGTTTATCTCGCGAAGGTGAGAAAAATTC 300
DB 1731 CTGCGGTATATACCGGTATTTGAGAGCAAGTTTATCTCGCGAAGGTGAGAAAAATTC 1790
QY 301 CTTCGCGAAGACGCTTGGCTGCAAGACGCTTTCAGCCGGAATACCGCTTTATTCGCGGC 360
DB 1791 CTTCGCGAAGACGCTTGGCTGCAAGACGCTTTCAGCCGGAATACCGCTTTATTCGCGGC 1850
QY 361 TTGGAAACGATGTTTATAGACGCTTCGACCTCGCCCTCGGCGGTGCGGATTAATGCGGG 420
DB 1851 TTGGAAACGATGTTTATAGACGCTTCGACCTCGCCCTCGGCGGTGCGGATTAATGCGGG 1910
QY 421 CGGCGCTTTCGCTGTTTGAAGAGCAACTGCGGCGGACGCGGCGCTATATCTCCGA 480
DB 1911 CGGCGCTTTCGCTGTTTGAAGAGCAACTGCGGCGGACGCGGCGCTATATCTCCGA 1970
QY 481 AAAGCGATGCGGTTTCTTGAAGAGGTTTGCGCCCTCGCCCGGACGCGCTGCAACCC 540
DB 1971 AAAGCGATGCGGTTTCTTGAAGAGGTTTGCGCCCTCGCCCGGACGCGCTGCAACCC 2030
```

```

RESULT 3
US-10-467-657-1309
; Sequence 1309, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467.657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMan99, version 1.04
; SEQ ID NO 1309
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1309

```

OY	361	TTGGAAAGATGTTTATATGACGTCGTGACCTCGCCCTCGGGGTGGCGGATTACTGGGG	420
Db	361	TTGGAAAGATGTTTATATGACGTCGTGACCTCGCCCTCGGGGTGGCGGATTACTGGGG	420
OY	421	CGCGCTTTCGCTGTGTGAAAGCGAACTGGGGGACGGCGGGCTATATCATTTTCCGA	480
Db	421	CGCGCTTTCGCGTTTGGAAAGCGAACACTGGCGGGACGGCGGGGTATATTTTCCGA	480
OY	481	AAAGCGATGCGGTTTTTCTGTGACAGTTTTGCGCGCTGCGCCCGAAGGGCTGCACCC	540
Db	481	AAAGGGAATGCTTTTTTCTGTGACAGTTTTGCGCTTTCGCCCGAAGCCTGCACCC	540
OY	541	GTGCACTGATGATGTGTACGGAATTTTTCGACAGGGAAGGAATGCGGTTTGCCAGCTC	600
Db	541	GTGCAATTTGATGATGTGTGGCACTCTGACAGAGGAAGGAATGCGGTTTGGCACTC	600
OY	601	AATCCCGCTTGTGTGCGCCCAAGAGCTGTGATTATGCGAATTTTCAAGCCAAACAGCGCA	660
Db	601	AATCCCGCTTGTGTGCGCCCAAGAGCTGTGATTATGCGAATTTTCAAGCCAAACAGCGCA	660
OY	661	TTTGGCAGCTGATGTGAACAGACCGGCTCTGTGAACCGGAACAGGAAGGGCGCATTC	720
Db	661	TTTGGCAGCTGATGTGAACATGACCGCGCTGGAACCGGAACAGGAATGGCGCATTC	720
OY	721	CCCGCAACATTTCAAAACACCGCTGTATCGCGCTTGACCAAAATCGACGAGGAAAG	780
Db	721	CCCGCAACATTTCAAAACACCGCTGTATCGCGCTTGACCAAAATCGGAGGAAAG	780
OY	781	GAAGAAAGCGCGCAAGGCGCGAAGCTTCATG	814
Db	781	GAAGAAAGCGCGCAAGGCGCGAAGCTTC	814

QY	DB	Query Match	Best Local Similarity	Matches	551, Conservative	57.0%, Score 471.8, DB 8, Length 840;	80.7%, Pred. No. 2.2e-136;	0, Mismatches 132;	Indels	0, Gaps	0
QY	Db	1	ATGCAAAACGACGTTATCAGCTTAGCTCCGCGCGAGAACGACGAGGGCGCATTTGCCGAT	60							
QY	Db	1	ATGCAAAACGACGTTATCAGCTTAGCTCCGCGCGAGAACGACGAGGGCGCATTTGCCGCA	60							
QY	Db	61	ACCTTCGGCAGGACGCGGATCCCGTTTCAGTTTTTCGACGACCTGATGCGCGCTCTGAAAG	120							
QY	Db	61	ACCTTCGGCAGGACGCGGATCCCGTTTCAGTTTTTCGACGACCTGATGCGCGCTCTGAAAG	120							
QY	Db	121	CTGGAACGAGGCAATGCGCGCAACTGCTCCCGGCGCTTGTGCGGCGACACCCCTATTGAGCGGA	180							
QY	Db	121	CTGGAACGAGGCAATGCGCGCAACTGCTCCCGGCGCTTGTGCGGCGACACCCCTATTGAGCGGA	180							

Qy	181	GTGAAAAAGCTGTTTATATAGCAGCCGTAATTGTGAAGCAAGCATTTGAGCAAGGT	240
Db	181	GTGAAAAAGCTGTTTATATAGCAGCCGTAATTGTGAAGCAAGCATTTGAGCAAGGT	240
Qy	241	CTGCAGTATACACGATTTTGAAGACAGCTTTACTCGCGAAAGTGAAGAAAAATTC	300
Db	241	TTACCGTATATTCGCGTATTTGAGATATGTTCTGCTTGGCAAGACGCAAAAAGTTT	300
Qy	301	CTTGCAGAAACGCTTGCTCAAGAACGCTTTGACCCGATTAACCGCTTTATGTCGCG	360
Db	301	CTTGCAGAAATACCTTGCTTGAAGAGGCTTTGATTAAGATATTCGCTTTATGTCGCGT	360
Qy	361	TTGGAACAGATGTTTATCAGCTCTGACCTCGCCCTCGGCGAGGCGAGTTATCGCGGG	420
Db	361	TTGGAACAGATGTTTGCAGAAAGTTATGTCAAGCCGATTAAGTCTGATTTAGAAAC	420
Qy	421	CGCGCTTTCCGCTGTGGAAGAAACGACATCGGGGAGAGGGGCTATATCATTTCCCGA	480
Db	421	CGGTCATTTCTTGTCTGAGAGCGACATGTGTGAAGCGGCTGGCTATATCATTTCCGCT	480
Qy	481	AAAGCATGCGGTTTTCTGTGAACAGGTTTCCCGCCCTCGCGCCGGAAGGCTGACCCC	540
Db	481	GAGCGAATGCGGTTTTCTGTGACAGGTTTCCGTTTTGCGCGACAGCCGATTTAAAGCG	540
Qy	541	GTCGATCTGATGATGTTCAAGGATTTTTTTCACAGGGAAGGATTCGCGTTTGGCAGCTC	600
Db	541	GTAGATTTGATGATGTTTACTTATTTCTTTGATAAGAGGGGATGCTGTTTATCAAGTT	600
Qy	601	AATCCGCGCTGTGCGCCCAAGAGCTGATATATGCAAGTTTTCACAGCAAAAACGCGCA	660
Db	601	AGTCCGCGCTTATGTACCCAGAAATTGCAATTATGCGAAATTTCTCACTCAAAACGATGTG	660
Qy	661	TTGGGACGCTGATCGAAACACGA 683	
Db	661	TTGGGTACCGATTTGGAAAAAGA 683	

```

RESULT 5
US-11-194-246-26
Sequence 26, Application US/11194246
Publication No. US20050272089A1
GENERAL INFORMATION:
APPLICANT: Mott, John
APPLICANT: Trepod, Catherine
APPLICANT: Arvidsson, Staffan
TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND ME
TITLE OF INVENTION: USE
FILE REFERENCE: 00592.US1 (MAR 268.05920101)
CURRENT APPLICATION NUMBER: US/11/194,246
CURRENT FILING DATE: 2005-08-01
PRIOR APPLICATION NUMBER: US/10/274,586
PRIOR FILING DATE: 2002-10-21
PRIOR APPLICATION NUMBER: US 60/345,438
PRIOR FILING DATE: 2001-10-19
NUMBER OF SEQ ID NOS: 621
SOFTWARE: PatentIn version 3.0
SEQ ID NO 26
LENGTH: 2449
TYPE: DNA
ORGANISM: ARTIFICIAL
FEATURE:
OTHER INFORMATION: Sequence of the ribB coding sequence and flanking regions.
US-11-194-246-26

```

Db	1632	ATTGCGTAGCGAATATTCCTTCGATTTTTCGATGCTATTAGCCCGCATTATTGA	1691
Oy	123	GGAAcAGGc--AATGGCGGAACTGTCGCCGCTTGCGCGCAcCCCTTATTGAGCGG	179
Db	1692	AGAAACCGCTAAAAAATTTATATATTACATTAGATCGCTCTCTTAAGCGAAAGTTGTCGA	1751
Oy	180	AGTGGAAAAAGCCTGCTTTATGAGCCAGCCCGTATTGGAAGCAGGCAATTGACGAGG	239
Db	1752	TGGGGAATGGTGTGCATTAACCAATATGTTTATGGCATTTAGCAATAGAAAAATAA	1811
Oy	240	TCTCCGCTATTCACCGTATTGAGAGAGAGTTTTACTCGGCGCAAGGTGAGAAAAATT	299
Db	1812	TTTAACTATATCAATATCTTTGAAGATGATATTCATTTGGGGGAAAAATGCCAAAGATT	1871
Oy	300	CCTTG 304	
Db	1872	ATTAG 1876	

```

RESULT 6
US-11-024-959-37
Sequence 37, Application US/11024959
Publication No. US20060010516a1
GENERAL INFORMATION:
APPLICANT: FORSTER, RICHARD L.
APPLICANT: CONNETT, MARIE B.
APPLICANT: EMERSON, SARAH JANE
APPLICANT: GRIGOR, MORRAY ROBERT
APPLICANT: HIGGINS, COLLEEN M.
APPLICANT: LUND, STEVEN TROY
APPLICANT: MAGUSIN, ANDREAS
APPLICANT: KORZYCKI, BOB
TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
FILE REFERENCE: 044463-0360
CURRENT APPLICATION NUMBER: US/11/024, 959
CURRENT FILING DATE: 2004-12-30
PRIOR APPLICATION NUMBER: 60/533, 036
PRIOR FILING DATE: 2003-12-30
NUMBER OF SEQ ID NOS: 782
SOFTWARE: PatentIn version 3.3
SEQ ID NO 37
LENGTH: 1760
TYPE: DNA
ORGANISM: Eucalyptus sp.
US-11-024-959-37

```

```

Query Match      4.6% ; Score 38 ; DB 14 ; Length 1760 ;
Query Local Similarity 50.5% ; Pred. No. 0.28 ;
Matches 92 ; Conservative 0 ; Mismatches 90 ; Indels 0 ; Gaps 0 ;

QY      590  TTTGCGAGCTCAATCCGCGCTTGCGGCCCAAGAGCTGATATATGCCAAGTTTTCAGACC 649
Db      123  TCTACGAGCCGACGATCGGCGACTACTATCTACGCGGACAGGGGCACTCGATGAAGCCCCACA 182

QY      650  AAAACACAGCGATTGGGAGCGCTGATTCGACACGACCGGCTCTTGAACCGGAAACAGCAA 709
Db      183  GGATCCGAGATGGCGGACAACTCTATCGTCCACTACCTTCACCGCGCATGGAGATCA 242

QY      710  GCGCGATTCGCCCCCGCCCAACATTTCAAACACGCGCTGATTCGCGCGCTTGCACAAAATCA 769
Db      243  GCGCGCGCTTCCCGCGCGCCACGACGACATCGCGGCTTCACTCCGAGGACTACGTCA 302

QY      770  GC 771
Db      303  CC 304

RESULT 7
US-10-997-437A-9/c
; Sequence 9, Application US/10997437A
; Publication No. US20050250184A1
; GENERAL INFORMATION:

```



```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 826
; LENGTH: 118864
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) (118864)
; OTHER INFORMATION: n = A,T,C or G
US-10-330-773-826
```

```
Query Match
Best Local Similarity 45.6%; Score 33.8; DB 9; Length 118864;
Pred. No. 20;
Matches 77; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
```

```
QY 611 TGTGCGCCGAGTGTGATTTAGCCAGTTTCACGACCAAAAGCGGATGGGACGCC 670
DB 71516 TGTGTGACAGAAATTGGGACCTGTGGAGTCTCTAGAGNNNNNNNNNNNNNNNNNN 71457
QY 671 TGAATGAACACGACCGCTCTCTGAACTCCGAAACAGAGGCGGATTTCCCGCCACA 730
DB 71456 TCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 71397
QY 731 CATTCAACACCGCGCTGATCCGCGCTTGACCAAAATCAGAGGGAAG 779
DB 71396 CAGCAGCAGAGCAGCATGATCTGTGTTTATTCATTAGGAAG 71348
```

RESULT 11

```
US-09-925-065A-5278#
; Sequence 52784, Application US/0925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52784
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-52784
```

```
Query Match
Best Local Similarity 52.9%; Score 33.6; DB 6; Length 535;
Pred. No. 4.6;
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```
QY 353 TGTCTCGCTTGAAGAAAGATTTATGACGTCCTGACCTTCCGCGGTGGCGGATT 412
DB 131 TGTGTGAGACTGGACACCCCTAGTTGTACCTGCTCTCTCTCTCTCTCTCTCTCTCT 190
QY 413 ACTGCGGGGCGCGCTTTCGCTGTGAAAGCGAAGCACTGGGGAGAGCGCGGCTATATCA 472
DB 191 TGAAGAGGACCTCTGGCCACACAGGAGGGCGACCCCAAGCTGCGGGTCTCCGATACCA 250
QY 473 TTTCGCGAAAAGGAT 488
DB 251 GATGCTGGAATGTGTT 266
```

```
RESULT 12
US-10-301-480-154022
; Sequence 154022, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 154022
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-154022
```

```
Query Match
Best Local Similarity 52.9%; Score 33.6; DB 9; Length 535;
Pred. No. 4.6;
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```
QY 353 TGTCTCGCTTGAAGAAAGATTTATGACGTCCTGACCTTCCGCGGTGGCGGATT 412
DB 131 TGTGTGAGACTGGACACCCCTAGTTGTACCTGCTCTCTCTCTCTCTCTCTCTCTCT 190
QY 413 ACTGCGGGGCGCGCTTTCGCTGTGAAAGCGAAGCACTGGGGAGAGCGCGGCTATATCA 472
DB 191 TGAAGAGGACCTCTGGCCACACAGGAGGGCGACCCCAAGCTGCGGGTCTCCGATACCA 250
QY 473 TTTCGCGAAAAGGAT 488
DB 251 GATGCTGGAATGTGTT 266
```

RESULT 13

```
US-10-301-480-767431
; Sequence 767431, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 767431
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-767431
```

```
Query Match
Best Local Similarity 52.9%; Score 33.6; DB 10; Length 535;
Pred. No. 4.6;
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```
QY 353 TGTCTCGCTTGAAGAAAGATTTATGACGTCCTGACCTTCCGCGGTGGCGGATT 412
DB 131 TGTGTGAGACTGGACACCCCTAGTTGTACCTGCTCTCTCTCTCTCTCTCTCTCTCT 190
QY 413 ACTGCGGGGCGCGCTTTCGCTGTGAAAGCGAAGCACTGGGGAGAGCGCGGCTATATCA 472
```

Db 191 TCAGAGGCGACCTCTGGCCACACAGGCGGCGACCCCGGCTCTCCGATACCA 250
QY 473 TTTCCGAAAAGGAT 488
Db 251 GATGCTGGAATGTGT 266

RESULT 14

US-11-096-281-10/c
; Sequence 10, Application US/11096281
; Publication No. US20050267036A1
; GENERAL INFORMATION:
; APPLICANT: GARRY, MARY
; APPLICANT: BEZPROZVANNY, ILYA
; TITLE OF INVENTION: PEPTIDES OF CAV2.2 THAT INHIBIT PAIN
; FILE REFERENCE: UTSD:1558US
; CURRENT APPLICATION NUMBER: US/11/096,281
; PRIOR FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: 60/558,383
; PRIOR FILING DATE: 2004-04-01
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 7364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (146)..(7165)
US-11-096-281-10

Query Match

Best Local Similarity 4.1%; Score 33.6; DB 14; Length 7364;
Matches 51; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 255 CGTATTGAGAGCAGCTTCTCGGCGAAGGTGAGAAAATCTTGGCCGAGAGCG 314
Db 407 CGTATTGAGAGCAGCTTCTCGGCGAAGGTGAGAAAATCTTGGCCGAGAGCG 348
QY 315 TTGGCTGCAAGAACGCTTG 334
Db 347 AGTTCTGCTTGACCGGAGTG 328

RESULT 15

US-09-925-065A-525973
; Sequence 525973, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 525973
; LENGTH: 626
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-525973

Query Match 4.0%; Score 33.4; DB 6; Length 626;
Best Local Similarity 47.8%; Pred. No. 5.6; Mismatches 106; Indels 0; Gaps 0;
Matches 97; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 344 CCGCCCTTATTCGTCGCTTGAAGCATGTTTATGACGCTCTGACCTCGCCCTCGGCG 403
Db 26 CCGCTTTACAGCCCTCAGAGAGATGACCCCATCCACACTTGATCTTGACTTTGGC 85
QY 404 TGGCGGATTACTGCGGCGCGCTTTCGCTGTTGAAAAGCAACTGGGGAGCGCG 463
Db 86 CTCGAGAACTGAGAGACAGTAGCTTCTGTGTTCTAAGTCAACCCAGTGTGTACTAT 145
QY 464 GCTATATCATTTCCGAAAAGCATGCGGTTTCTCGACAGCTTTCGCCCTCGCGC 523
Db 146 GCTATGCGAGCCCTAGAAAATGAAATTCAGATTATCAACCTGTTGAGATGTCAGC 205
QY 524 CCGAAGGCTGCACCCCGTCGAT 546
Db 206 CCGTTGGCAGAGCCAGGAGCT 228

Search completed: April 7, 2006, 21:23:57
Job time : 969.604 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:20:42 ; Search time 9845.41 Seconds
(without alignments)
4650.028 Million cell updates/sec

Title: US-09-211-691-1

Sequence: 1 acgcaaacacgtatcag.....tcattgcttcacataa 828

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 79147668 seqs, 27645789525 residues

Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*

Result No.	Score	Query Match	Length	ID	Description
1	828	100.0	828	22	US-09-211-691-1
2	828	100.0	828	49	US-10-317-428-1
3	828	100.0	828	49	US-10-317-773-1
4	826.4	99.8	46594	33	US-09-806-866A-4
5	826.4	99.8	46594	63	US-10-915-740A-4
6	826.4	99.8	224716	33	US-09-806-866A-1068
7	826.4	99.8	224716	63	US-10-915-740A-1068
8	826.4	99.8	227325	40	US-10-018-470A-1
9	826.4	99.8	5859	40	US-10-096-129-1
10	726	87.7	5859	54	US-10-654-528-7
11	726	87.7	5859	54	US-10-654-528-7
12	726	87.7	5859	70	US-11-102-497-1
13	726	87.7	5859	70	US-11-102-497-1
14	721.2	87.1	837	51	US-10-467-657A-1309
15	721.2	87.1	837	51	US-10-467-657A-1309
16	703.4	85.0	9906	74	US-60-068-138-355
17	471.8	57.0	840	51	US-10-467-657A-1325
18	471.8	57.0	840	51	US-10-467-657A-1325
19	426	51.4	840	51	US-10-472-260-175
20	93	11.2	939	82	US-60-691-214-1378
21	93	11.2	14547	2	PCT-US03-17092-570
22	93	11.2	14547	2	PCT-US04-07001-570
23	93	11.2	14547	61	US-10-795-159-570

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES


```
QY 1 ATGCAAAACCAAGTTATGAGCTTCCGCGCAGAAAGGCGCAATTGCCGAT 60
DB 1 ATGCAAAACCAAGTTATGAGCTTCCGCGCAGAAAGGCGCGCAATTGCCGAT 60
QY 61 ACCCTGGCAGGCAAGCGGCAATCCGTTTCACTTTTTCAGCATGATCCGTCGAAAG 120
DB 61 ACCCTGGCAGGCAAGCGGCAATCCGTTTCACTTTTTCAGCATGATCCGTCGAAAG 120
QY 121 CTGGAACAGGCAATGCGGAACTGTCCTCCGCGCTTGCAGCGCAATTTTGAAGCGA 180
DB 121 CTGGAACAGGCAATGCGGAACTGTCCTCCGCGCTTGCAGCGCAATTTTGAAGCGA 180
QY 181 GTGGAAGAAAGCTGCTTTATGAGCAAGCGCTATTTGGAAGAGGCAATTTGAAGGT 240
DB 181 GTGGAAGAAAGCTGCTTTATGAGCAAGCGCTATTTGGAAGAGGCAATTTGAAGGT 240
QY 241 CTGCGCTATATCAACGTAATTTGAGAGCAAGTTTACTCGCGCAAGGTGAGAAATTC 300
DB 241 CTGCGCTATATCAACGTAATTTGAGAGCAAGTTTACTCGCGCAAGGTGAGAAATTC 300
QY 301 CTGCGCAAGAGCGCTTGGCTGCAAGAAAGCTTTGACCGCGATACCGCTTTATCGTCGC 360
DB 301 CTGCGCAAGAGCGCTTGGCTGCAAGAAAGCTTTGACCGCGATACCGCTTTATCGTCGC 360
QY 361 TTGGAACAGATGTTTATGACAGTCTGACCTCGCCCTCGCGGCGGATTAATGAGG 420
DB 361 TTGGAACAGATGTTTATGACAGTCTGACCTCGCCCTCGCGGCGGATTAATGAGG 420
QY 421 CGGCGCTTTCCGCTGTTGGAAGAGCACTGCGGAGGAGCGGCGCTATATCATTTCCGA 480
DB 421 CGGCGCTTTCCGCTGTTGGAAGAGCACTGCGGAGGAGCGGCGCTATATCATTTCCGA 480
QY 481 AAAGCATGCGGTTTCTCGACAGGTTGCGCGCTCGCGCGCGCAAGGCGTGCACCC 540
DB 481 AAAGCATGCGGTTTCTCGACAGGTTGCGCGCTCGCGCGCGCAAGGCGTGCACCC 540
QY 541 GTGCAATGATGATGTTTATGACAGATTTTTCAGAGGAGAGGATGCGGTTTCCAGCTC 600
DB 541 GTGCAATGATGATGTTTATGACAGATTTTTCAGAGGAGAGGATGCGGTTTCCAGCTC 600
QY 601 AATCCCGCTTGTGCGCCCAAGAGTGCATTATGCAAGTTTCAAGCAACCAACAGCGCA 660
DB 601 AATCCCGCTTGTGCGCCCAAGAGTGCATTATGCAAGTTTCAAGCAACCAACAGCGCA 660
QY 661 TTGGGAGAGCTGATTCGAACAGCAAGCTCTCTGCAACCGCAACAGCAAGGCGCATTC 720
DB 661 TTGGGAGAGCTGATTCGAACAGCAAGCTCTCTGCAACCGCAACAGCAAGGCGCATTC 720
QY 721 CCGGCAACAGATTTCAAAACAGCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 780
DB 721 CCGGCAACAGATTTCAAAACAGCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 780
QY 781 GAAAAAGCGCGGCAAGGCGGCAAGGTCATTGTGCTTTTCCAAATTA 828
DB 781 GAAAAAGCGCGGCAAGGCGGCAAGGTCATTGTGCTTTTCCAAATTA 828

RESULT 3
US-10-317-773-1
; Sequence 1, Application US/10317773
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Fusion Protein Comprising a UDP-GalNAc 4' Epimerase and a
; FILE REFERENCE: 019633-000812US
; CURRENT APPLICATION NUMBER: US/10/317,773
; CURRENT FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
```

```
; PRIOR APPLICATION NUMBER: US 09/211,691
; PRIOR FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(828)
; OTHER INFORMATION: beta-1,4-galactosyltransferase (19cB)
US-10-317-773-1

Query Match 100.0%; Score 828; DB 49; Length 828;
Best Local Similarity 100.0%; Pred. No. 5,6e-282;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAAAACCAAGTTATGAGCTTCCGCGCAGAAAGGCGCAATTGCCGAT 60
DB 1 ATGCAAAACCAAGTTATGAGCTTCCGCGCAGAAAGGCGCGCAATTGCCGAT 60
QY 61 ACCCTGGCAGGCAAGCGGCAATCCGTTTCACTTTTTCAGCATGATCCGTCGAAAG 120
DB 61 ACCCTGGCAGGCAAGCGGCAATCCGTTTCACTTTTTCAGCATGATCCGTCGAAAG 120
QY 121 CTGGAACAGGCAATGCGGAACTGTCCTCCGCGCTTGCAGCGCAATTTTGAAGCGA 180
DB 121 CTGGAACAGGCAATGCGGAACTGTCCTCCGCGCTTGCAGCGCAATTTTGAAGCGA 180
QY 181 GTGGAAGAAAGCTGCTTTATGAGCAAGCGCTATTTGGAAGAGGCAATTTGAAGGT 240
DB 181 GTGGAAGAAAGCTGCTTTATGAGCAAGCGCTATTTGGAAGAGGCAATTTGAAGGT 240
QY 241 CTGCGCTATATCAACGTAATTTGAGAGCAAGTTTACTCGCGCAAGGTGAGAAATTC 300
DB 241 CTGCGCTATATCAACGTAATTTGAGAGCAAGTTTACTCGCGCAAGGTGAGAAATTC 300
QY 301 CTGCGCAAGAGCGCTTGGCTGCAAGAAAGCTTTGACCGCGATACCGCTTTATCGTCGC 360
DB 301 CTGCGCAAGAGCGCTTGGCTGCAAGAAAGCTTTGACCGCGATACCGCTTTATCGTCGC 360
QY 361 TTGGAACAGATGTTTATGACAGTCTGACCTCGCCCTCGCGGCGGATTAATGAGG 420
DB 361 TTGGAACAGATGTTTATGACAGTCTGACCTCGCCCTCGCGGCGGATTAATGAGG 420
QY 421 CGGCGCTTTCCGCTGTTGGAAGAGCACTGCGGAGGAGCGGCGCTATATCATTTCCGA 480
DB 421 CGGCGCTTTCCGCTGTTGGAAGAGCACTGCGGAGGAGCGGCGCTATATCATTTCCGA 480
QY 481 AAAGCATGCGGTTTCTCGACAGGTTGCGCGCTCGCGCGCGCAAGGCGTGCACCC 540
DB 481 AAAGCATGCGGTTTCTCGACAGGTTGCGCGCTCGCGCGCGCAAGGCGTGCACCC 540
QY 541 GTGCAATGATGATGTTTATGACAGATTTTTCAGAGGAGAGGATGCGGTTTCCAGCTC 600
DB 541 GTGCAATGATGATGTTTATGACAGATTTTTCAGAGGAGAGGATGCGGTTTCCAGCTC 600
QY 601 AATCCCGCTTGTGCGCCCAAGAGTGCATTATGCAAGTTTCAAGCAACCAACAGCGCA 660
DB 601 AATCCCGCTTGTGCGCCCAAGAGTGCATTATGCAAGTTTCAAGCAACCAACAGCGCA 660
QY 661 TTGGGAGAGCTGATTCGAACAGCAAGCTCTCTGCAACCGCAACAGCAAGGCGCATTC 720
DB 661 TTGGGAGAGCTGATTCGAACAGCAAGCTCTCTGCAACCGCAACAGCAAGGCGCATTC 720
QY 721 CCGGCAACAGATTTCAAAACAGCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 780
DB 721 CCGGCAACAGATTTCAAAACAGCGCTGATCCGCGCTTGAACCAAAATCAGCAGGAAAG 780
QY 781 GAAAAAGCGCGGCAAGGCGGCAAGGTCATTGTGCTTTTCCAAATTA 828
DB 781 GAAAAAGCGCGGCAAGGCGGCAAGGTCATTGTGCTTTTCCAAATTA 828
```

```

RESULT 4
US-09-806-866A-4/c
; Sequence 4, Application US/09806866A
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tetteilin, Herre
; APPLICANT: Venter, J. Craig
; APPLICANT: Massignani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroza
; APPLICANT: Ratti, Giulio
; APPLICANT: Scarfello, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: CHIR0313
; CURRENT APPLICATION NUMBER: US/09/806,866A
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: USSN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USSN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 46594
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-09-806-866A-4

Query Match          99.8%; Score 826.4; DB 33; Length 46594;
Best Local Similarity 99.9%; Pred. No. 6.6e-280;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ATGCAAAACACGATTATGCTTACGCTCCGCGGCAAGCAAGGCGGCAATGGCCAT 60
DB      33537 ATGCAAAACACGATTATGCTTACGCTCCGCGGCAAGCAAGGCGGCAATGGCCAT 33478
QY      61 ACCTTCGAGGACGCGCATCCCGTTTCAGTTTTCAGCACTGATGCGCTGTAAGG 120
DB      33477 ACCTTCGAGGACGCGCATCCCGTTTCAGTTTTCAGCACTGATGCGCTGTAAGG 33418
QY      121 CTGGAACAGGCAATGGCGGAACCTCGCCCGGCTGTGGCCGACCCCTATTGAGCGGA 180
DB      33417 CTGGAACAGGCAATGGCGGAACCTCGCCCGGCTGTGGCCGACCCCTATTGAGCGGA 33358
QY      181 GTGAAAAAAGCTGCTTTATGAGCCAGCGCGTATTGTGAGACAGGCAATTGACGAAGGT 240
DB      33357 GTGAAAAAAGCTGCTTTATGAGCCAGCGCGTATTGTGAGACAGGCAATTGACGAAGGT 33298
QY      241 CTGCGGTATATACGCGTATTGAGAGCGAGCTTTTACTGCGGGAAGGTGAGAAAAATTC 300
DB      33297 CTGCGGTATATACGCGTATTGAGAGCGAGCTTTTACTGCGGGAAGGTGAGAAAAATTC 33238
QY      301 CTTCGCGAAGACGCTTGGCTGCAAGAACGCTTTGACCCGGAATACCGCCCTTTATCGTCCGC 360
DB      33237 CTTCGCGAAGACGCTTGGCTGCAAGAACGCTTTGACCCGGAATACCGCCCTTTATCGTCCGC 33178
QY      361 TTGGAAGAGATGTTTATGACAGCTCGACCTGCGCCCTCGGGGTGGCGGATTAATACGCGGG 420
DB      33177 TTGGAAGAGATGTTTATGACAGCTCGACCTGCGCCCTCGGGGTGGCGGATTAATACGCGGG 33118
QY      421 CGCGCTTTTCGCTGTTGAAAAAGCAACTGAGGGAGCGCGCGCTATATCATTTTCCGGA 480
DB      33117 CGCGCTTTTCGCTGTTGAAAAAGCAACTGAGGGAGCGCGCGCTATATCATTTTCCGGA 33058

```

```

QY      481 AAAGCATGCGGTTTTTCTTGACAGGTTTGCCGCGCTCCGCGCCGGAAGGCTGACCC 540
DB      33057 AAAGCATGCGGTTTTTCTTGACAGGTTTGCCGCGCTCCGCGCCGGAAGGCTGACCC 32998
QY      541 GTGCATCTGATGATGTTTACGAGATTTTTCGACAGGAAGAAATGCGGTTTGCAGCTC 600
DB      32997 GTGCATCTGATGATGTTTACGAGATTTTTCGACAGGAAGAAATGCGGTTTGCAGCTC 32938
QY      601 AATCCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTACGACCAAAACAGCGCA 660
DB      32937 AATCCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTACGACCAAAACAGCGCA 32878
QY      661 TTGGCAGCGTATGACACAGACCGCTCCGTCGAACCGCAACAGCAAGGCGCATTC 720
DB      32877 TTGGCAGCGTATGACACAGACCGCTCCGTCGAACCGCAACAGCAAGGCGCATTC 32818
QY      721 CCCGCAACACATTCAAAACCCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG 780
DB      32817 CCCGCAACACATTCAAAACCCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG 32758
QY      781 GAAAAACCGCGCAAGGCGCAACAGTTCATTGTCCTTTCATTA 828
DB      32757 GAAAAACCGCGCAAGGCGCAACAGTTCATTGTCCTTTCATTA 32710

RESULT 5
US-10-915-740A-4/c
; Sequence 4, Application US/10915740A
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tetteilin, Herre
; APPLICANT: Venter, J. Craig
; APPLICANT: Massignani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroza
; APPLICANT: Ratti, Giulio
; APPLICANT: Scarfello, Maria
; APPLICANT: Scarfello, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: 002441.00090
; CURRENT APPLICATION NUMBER: US/10/915,740A
; CURRENT FILING DATE: 2004-08-11
; PRIOR APPLICATION NUMBER: 09/806,866
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: USSN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USSN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 46594
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-915-740A-4

Query Match          99.8%; Score 826.4; DB 63; Length 46594;
Best Local Similarity 99.9%; Pred. No. 6.6e-280;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ATGCAAAACACGATTATGCTTACGCTCCGCGGCAAGCAAGGCGGCAATGGCCAT 60
DB      33537 ATGCAAAACACGATTATGCTTACGCTCCGCGGCAAGCAAGGCGGCAATGGCCAT 33478
QY      61 ACCTTCGAGGACGCGCATCCCGTTTCAGTTTTCAGACGCACTGATGCGCTGTAAGG 120

```

```

Db 33477 ACCTCGGCAGGACCGGCATCCGTTTCAGTTTTCAGACGATGCGTCAAGAG 33418
QY 121 CTGGAAACAGGCAATGGGGAACTGTCGCCGGCTGTGCGCGACCCCTTTTGAAGGA 180
Db 33417 CTGGAAACAGGCAATGGGGAACTGTCGCCGGCTGTGCGCGACCCCTTTTGAAGGA 33358
QY 181 GTGGAAAAAGCCCTGTTTATGACGACGCGCTATTGTGAGAGAGGATTTGAAGAGT 240
Db 33357 GTGGAAAAAGCCCTGTTTATGAGCGACGCGCTATTGTGAGAGAGGATTTGAAGAGT 33298
QY 241 CTGCGCATATATCAACCGTATTGAGAGAGCGTTTACTGCGCGAAGGTGAGAAAAATTC 300
Db 33297 CTGCGCATATATCAACCGTATTGAGAGAGCGTTTACTGCGCGAAGGTGAGAAAAATTC 33238
QY 301 CTGGCGGAAGACGCTTGTGCGCAAGAACGCTTTGACCCCGATACCGCTTTATGTCGCC 360
Db 33237 CTGGCGGAAGACGCTTGTGCGCAAGAACGCTTTGACCCCGATACCGCTTTATGTCGCC 33178
QY 361 TTGGAAACGATGTTTATGACAGTCTGACCTGCGCCCTCGCGCGAGATTACTGCGAG 420
Db 33177 TTGGAAACGATGTTTATGACAGTCTGACCTGCGCCCTCGCGCGAGATTACTGCGAG 33118
QY 421 CGGCGCTTTCCGCTGTTGAAAGCGAACACTGCGAGGAGCGCGCTATATCATTTCCGA 480
Db 33117 CGGCGCTTTCCGCTGTTGAAAGCGAACACTGCGAGGAGCGCGCTATATCATTTCCGA 33058
QY 481 AAAGCAGTGGGTTTTCCTGCGACAGTTTGGCGCCCTGCGCGCGAAGGGCTGCAAGG 540
Db 33057 AAAGCAGTGGGTTTTCCTGCGACAGTTTGGCGCCCTGCGCGCGAAGGGCTGCAAGG 32998
QY 541 GTGCACTGATGATGTTTTCAGCGATTTTTCAGACAGGAGAGAAATGCGGTTTCCAGCTC 600
Db 32997 GTGCACTGATGATGTTTTCAGCGATTTTTCAGACAGGAGAGAAATGCGGTTTCCAGCTC 32938
QY 601 AATCCGCGCTTGTGCGCCCGCAAGAGCTGATTTATGCCAAGTTTTCAGACCAAAACAGCGCA 660
Db 32937 AATCCGCGCTTGTGCGCCCGCAAGAGCTGATTTATGCCAAGTTTTCAGACCAAAACAGCGCA 32878
QY 661 TTGGGAGAGCTGATTCGACACGACCGCTCTCTGACCGCAACAGCAAGGCGGATTC 720
Db 32877 TTGGGAGAGCTGATTCGACACGACCGCTCTCTGACCGCAACAGCAAGGCGGATTC 32818
QY 721 CCGCGCAACATTTCAAAACCGCGCTGATCCGCGCTTGAACCAAAATCGACAGGAAAG 780
Db 32817 CCGCGCAACATTTCAAAACCGCGCTGATCCGCGCTTGAACCAAAATCGACAGGAAAG 32758
QY 781 GAAAAACGCGCGCAAGGCGCGAACAGTTCAATTGCTTTCATTA 828
Db 32757 GAAAAACGCGCGCAAGGCGCGAACAGTTCAATTGCTTTCATTA 32710

```

```

RESULT 6
US-09-806-866A-1068/c
; Sequence 1068, Application US/09806866A
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tetteilin, Herve
; APPLICANT: Venter, J. Craig
; APPLICANT: Masimani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Manroza
; APPLICANT: Rattai, Giulio
; APPLICANT: Scarselli, Maria
; APPLICANT: Scariato, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: CH10313
; CURRENT APPLICATION NUMBER: US/09/806,866A

```

```

; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: USN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1068
; LENGTH: 2242716
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-09-806-866A-1068

Query Match 99.8%; Score 826.4; DB 33; Length 2242716;
Best Local Similarity 99.9%; Pred. No. 1.9e-278;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCAAAACAGGTTATGAGCTTAGCTTCGCGCGCAAGACGAGGCGGACATTTGCGCAT 60
Db 1997047 ATGCAAAACAGGTTATGAGCTTAGCTTCGCGCGCAAGACGAGGCGGACATTTGCGCAT 1996988
QY 61 ACCTTGCGAGGACGCGCATCCCGTTTCACTTTTTCAGAGCACTGATGCGCTGAAAG 120
Db 1996987 ACCTTGCGAGGACGCGCATCCCGTTTCACTTTTTCAGAGCACTGATGCGCTGAAAG 1996928
QY 121 CTGGAAACGCAATGGCGGAACCTGTCGCCGCTGTGCGCGACCCCTATTGAGCGGA 180
Db 1996927 CTGGAAACGCAATGGCGGAACCTGTCGCCGCTGTGCGCGACCCCTATTGAGCGGA 1996868
QY 181 GTGGAAAAAGCCCTGTTTATGAGCGACGCGCGATTTGTGAGAGAGGATTTGAAGAGT 240
Db 1996867 GTGGAAAAAGCCCTGTTTATGAGCGACGCGCGATTTGTGAGAGAGGATTTGAAGAGT 1996808
QY 241 CTGCGCATATATCAACCGTATTGAGAGAGAGTTTTCAGCGCAAGGTGAGAAAAATTC 300
Db 1996807 CTGCGCATATATCAACCGTATTGAGAGAGAGTTTTCAGCGCAAGGTGAGAAAAATTC 1996748
QY 301 CTGGCGGAAGAGCTTGTGCGCAAGAGCTTGAACCGCGGATTAACCGCTTTATGTCGCC 360
Db 1996747 CTGGCGGAAGAGCTTGTGCGCAAGAGCTTGAACCGCGGATTAACCGCTTTATGTCGCC 1996688
QY 361 TTGGAAACGATGTTTATGACAGTCTGACCTGCGCTCGCGGTGCGGATTAATGCGGG 420
Db 1996687 TTGGAAACGATGTTTATGACAGTCTGACCTGCGCTCGCGGTGCGGATTAATGCGGG 1996628
QY 421 CCGCGCTTTCCGCTGTTGAAAGCGAACACTGAGGAGAGCGAGGCTATATCATTTCCGA 480
Db 1996627 CCGCGCTTTCCGCTGTTGAAAGCGAACACTGAGGAGAGCGAGGCTATATCATTTCCGA 1996568
QY 481 AAAGCAGTGGGTTTTCCTGGAACAGTTTTCGCGCTGCGCGCGCAAGGCGTCAAGG 540
Db 1996567 AAAGCAGTGGGTTTTCCTGGAACAGTTTTCGCGCTGCGCGCGCAAGGCGTCAAGG 1996508
QY 541 GTGCACTGATGATGTTTTCAGCGATTTTTCAGACAGGAGAGAAATGCGGTTTCCAGCTC 600
Db 1996507 GTGCACTGATGATGTTTTCAGCGATTTTTCAGACAGGAGAGAAATGCGGTTTCCAGCTC 1996448
QY 601 AATCCGCGCTTGTGCGCCCGCAAGAGCTGATTTATGCAAGTTTTCAGACCAAAACAGCGCA 660
Db 1996447 AATCCGCGCTTGTGCGCCCGCAAGAGCTGATTTATGCAAGTTTTCAGACCAAAACAGCGCA 1996388
QY 661 TTGGGAGAGCTGATTCGACACGACCGCTCTCTGAAACCGCAACAGCAAGGCGGATTC 720
Db 1996387 TTGGGAGAGCTGATTCGACACGACCGCTCTCTGAAACCGCAACAGCAAGGCGGATTC 1996328
QY 721 CCGCGCAACATTTCAAAACCGCGCTGATCCGCGCTTGAACCAAAATCGACAGGAAAG 780
Db 1996327 CCGCGCAACATTTCAAAACCGCGCTGATCCGCGCTTGAACCAAAATCGACAGGAAAG 1996268
QY 781 GAAAAACGCGCGCAAGGCGCGAACAGTTCAATTGCTTTCATTA 828

```

Db 1996267 GAAAAACGCCGCAAGGCGGCAACAGTTCATTGCTTTCCATATA 1996220

RESULT 7

US-10-915-740A-1068/c
 ; Sequence 1068, Application US/10915740A
 ; GENERAL INFORMATION:
 ; APPLICANT: Frazer, Claire M.
 ; APPLICANT: Hickey, Erin
 ; APPLICANT: Peterson, Jeremy
 ; APPLICANT: Tettelin, Hervé
 ; APPLICANT: Venter, J. Craig
 ; APPLICANT: Massignani, Vega
 ; APPLICANT: Galeotti, Cesira
 ; APPLICANT: Mora, Manroa
 ; APPLICANT: Ratti, Giulio
 ; APPLICANT: Scarselli, Maria
 ; APPLICANT: Scarlatto, Vincenzo
 ; APPLICANT: Rappuoli, Rino
 ; APPLICANT: Pizza, Mariagrazia
 ; APPLICANT: Grandi, Guido
 ; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
 ; FILE REFERENCE: 002441.00090
 ; CURRENT APPLICATION NUMBER: US/10/915,740A
 ; PRIOR FILING DATE: 2004-08-11
 ; PRIOR APPLICATION NUMBER: 09/806,866
 ; PRIOR FILING DATE: 1999-10-08
 ; PRIOR APPLICATION NUMBER: USSN 60/103,794
 ; PRIOR FILING DATE: 1998-10-09
 ; PRIOR APPLICATION NUMBER: USSN 60/132,068
 ; PRIOR FILING DATE: 1999-04-30
 ; PRIOR APPLICATION NUMBER: PCT/US99/25373
 ; PRIOR FILING DATE: 1999-10-08
 ; NUMBER OF SEQ ID NOS: 1068
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO: 1068
 ; LENGTH: 2242716
 ; TYPE: DNA
 ; ORGANISM: Neisseria meningitidis
 ; US-10-915-740A-1068

Query Match 99.8%; Score 826.4; DB 63; Length 2242716;
 Best Local Similarity 99.9%; Pred. No. 1.9e-278;
 Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCAAAACCAAGTATGAGCTTACGCTTCCGCGCAAGACGAGCGGCAATTGCCGAT 60
 Db 1997047 ATGCAAAACCAAGTATGAGCTTACGCTTCCGCGCAAGACGAGCGGCAATTGCCGAT 1996988
 Qy 61 ACCTTGGGAGGAGCGGATCCCGTTTCAAGTTTTCAGAGCACTGATGCGTGAAGAAG 120
 Db 1996987 ACCTTGGGAGGAGCGGATCCCGTTTCAAGTTTTCAGAGCACTGATGCGTGAAGAAG 1996928
 Qy 121 CTGGAACAGCAATGCGGCAATCTGTCGCCGCTTGTGCGCGCACCCCTATTGAGCGGA 180
 Db 1996927 CTGGAACAGCAATGCGGCAATCTGTCGCCGCTTGTGCGCGCACCCCTATTGAGCGGA 1996868
 Qy 181 GTGGAAGAAAGCTGCTTATGAGCAGCGCGATTTGTGGAAGAGGCAATTGAGCAAGGT 240
 Db 1996867 GTGGAAGAAAGCTGCTTATGAGCAGCGCGATTTGTGGAAGAGGCAATTGAGCAAGGT 1996808
 Qy 241 CTGCGGTATATCAACGATTTTGAAGAGCAAGCTTTTACTGGGGAAGGTGAGAAAAATTTC 300
 Db 1996807 CTGCGGTATATCAACGATTTTGAAGAGCAAGCTTTTACTGGGGAAGGTGAGAAAAATTTC 1996748
 Qy 301 CTGCGCAAGAGCGCTTGTGCAAGAAAGCTTTGACCCGGAATTCGCTTTATCGTCCGC 360
 Db 1996747 CTGCGCAAGAGCGCTTGTGCAAGAAAGCTTTGACCCGGAATTCGCTTTATCGTCCGC 1996688
 Qy 361 TTGGAAGCAATGTTTATGACAGCTTCTGAGACTCGCCCTCGGCGGTGGCGGATTAATCGCGG 420
 Db 1996687 TTGGAAGCAATGTTTATGACAGCTTCTGAGACTCGCCCTCGGCGGTGGCGGATTAATCGCGG 1996628

Qy 421 CGCGCTTTCCGCTTTGGAAGCAACATGCGGAGGAGCGGCGGCTATATCATTTTCCGA 480
 Db 1996627 CGCGCTTTCCGCTTTGGAAGCAACATGCGGAGGAGCGGCGGCTATATCATTTTCCGA 1996568
 Qy 481 AAAGCATGCGGTTTCTTCTGACAGTTTTCGCCCTGCGGCCGGAAGGCTGCAACCC 540
 Db 1996567 AAAGCATGCGGTTTCTTCTGACAGTTTTCGCCCTGCGGCCGGAAGGCTGCAACCC 1996508
 Qy 541 GTGCATCTGATGATGTTTCAAGCATTTTTCGACAGGAGAAATCCGCTTTCGACGTC 600
 Db 1996507 GTGCATCTGATGATGTTTCAAGCATTTTTCGACAGGAGAAATCCGCTTTCGACGTC 1996448
 Qy 601 AATCCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTCAAGCAACCAAGAGCGCA 660
 Db 1996447 AATCCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTCAAGCAACCAAGAGCGCA 1996388
 Qy 661 TTGGGACGCTGATGGAACAGACCGCTCTGAAACCGCAACAGCAAGAGCGGATTC 720
 Db 1996387 TTGGGACGCTGATGGAACAGACCGCTCTGAAACCGCAACAGCAAGAGCGGATTC 1996328
 Qy 721 CCCGCCAACATTTGAAACACCGCTGATTCGCGCTTGAACCAAAATGACAGGAGAAAG 780
 Db 1996327 CCCGCCAACATTTGAAACACCGCTGATTCGCGCTTGAACCAAAATGACAGGAGAAAG 1996268
 Qy 781 GAAAAACGCCGCAAGGCGGCAACAGTTCATTGCTTTCCATATA 828
 Db 1996267 GAAAAACGCCGCAAGGCGGCAACAGTTCATTGCTTTCCATATA 1996220

RESULT 8

US-10-018-470A-1/c
 ; Sequence 1, Application US/10018470A
 ; GENERAL INFORMATION:
 ; APPLICANT: Frazer, Claire M.
 ; APPLICANT: Hickey, Erin
 ; APPLICANT: Peterson, Jeremy
 ; APPLICANT: Tettelin, Hervé
 ; APPLICANT: Venter, J. Craig
 ; APPLICANT: Massignani, Vega
 ; APPLICANT: Galeotti, Cesira
 ; APPLICANT: Mora, Manroa
 ; APPLICANT: Ratti, Giulio
 ; APPLICANT: Scarselli, Maria
 ; APPLICANT: Scarlatto, Vincenzo
 ; APPLICANT: Rappuoli, Rino
 ; APPLICANT: Pizza, Mariagrazia
 ; APPLICANT: Grandi, Guido
 ; TITLE OF INVENTION: Neisseria Genomic Research
 ; FILE REFERENCE: CHIR-0319
 ; CURRENT APPLICATION NUMBER: US/10/018,470A
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/132,068
 ; PRIOR FILING DATE: 1999-04-30
 ; PRIOR APPLICATION NUMBER: PCT/US99/25373
 ; PRIOR FILING DATE: 1999-10-08
 ; PRIOR APPLICATION NUMBER: GB-0004695.3
 ; PRIOR FILING DATE: 2000-02-28
 ; PRIOR APPLICATION NUMBER: PCT/US/05928
 ; PRIOR FILING DATE: 2000-03-08
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO: 1
 ; LENGTH: 2272325
 ; TYPE: DNA
 ; ORGANISM: Neisseria meningitidis
 ; US-10-018-470A-1

Query Match 99.8%; Score 826.4; DB 40; Length 2272325;
 Best Local Similarity 99.9%; Pred. No. 1.9e-278;
 Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCAAAACCAAGTATGAGCTTACGCTTCCGCGCAAGACGAGGCGCAATTGCCGAT 60
 |||

```
Db 202656 ATGCAAAACAGCTTATAGCTTCCGCGCAGAAAGGCGCATATGCGCAT 2026597
Qy 61 ACCCTGGGAGGAGGAGGATCCGTTCACTTTTGGACGATGATCCGCTGGAAG 120
Db 2026596 ACCCTGGGAGGAGGAGGATCCGTTCACTTTTGGACGATGATGCGCTGGAAG 2026537
Qy 121 CTGGAACAGGCAATGGCGGAATCTGTCGCCGCTTGTGCGGCAACCCCTATTGAGCGGA 180
Db 2026536 CTGGAACAGGCAATGGCGGAATCTGTCGCCGCTTGTGCGGCAACCCCTATTGAGCGGA 2026477
Qy 181 GTGAAAAAACCTGCTTTATGAGCAGCCGATTTGTGAAGCAGGATGAGCAAGGT 240
Db 2026476 GTGAAAAAACCTGCTTTATGAGCAGCCGATTTGTGAAGCAGGATGAGCAAGGT 2026417
Qy 241 CTGCGGATATATCAACGATTTTGAAGAGAGGTTTATCGGGCAAGGTGAGAAATTC 300
Db 2026416 CTGCGGATATATCAACGATTTTGAAGAGAGGTTTATCGGGCAAGGTGAGAAATTC 2026357
Qy 301 CTTCGGAAGAGCCTTGCTGCAAGACGCTTTGACCCGATACCGCTTTATCTGCGC 360
Db 2026356 CTTCGGAAGAGCCTTGCTGCAAGACGCTTTGACCCGATACCGCTTTATCTGCGC 2026297
Qy 361 TTGGAACAGATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGAGATTACTGCGG 420
Db 2026296 TTGGAACAGATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGAGATTACTGCGG 2026237
Qy 421 CGGCGCTTTCGCTGTTTGAAGAGCACTGGGGGAGCGCGGCTTATCATTTCCCGA 480
Db 2026236 CGGCGCTTTCGCTGTTTGAAGAGCACTGGGGGAGCGCGGCTTATCATTTCCCGA 2026177
Qy 481 AAGGAGATGCGTCTTCTGCAAGGTTGCGCGCTGCGCGCGCAAGGCTGCAACCC 540
Db 2026176 AAGGAGATGCGTCTTCTGCAAGGTTGCGCGCTGCGCGCGCAAGGCTGCAACCC 2026117
Qy 541 GTGATCTGATGATGTTTCAAGCAATTTTTCAGACAGGAGAAATGCCGTTTCCAGCTC 600
Db 2026116 GTGATCTGATGATGTTTCAAGCAATTTTTCAGACAGGAGAAATGCCGTTTCCAGCTC 2026057
Qy 601 AATCCGCGCTTGTGCGCGCAAGAGCTGATATGCAAGTTTACAGCAAAACAGCGCA 660
Db 2026056 AATCCGCGCTTGTGCGCGCAAGAGCTGATATGCAAGTTTACAGCAAAACAGCGCA 2025997
Qy 661 TTGGGAGGCTGATTCAGACAGACCGCTCTCTGAACCGCAAAAGGCGCGATTC 720
Db 2025996 TTGGGAGGCTGATTCAGACAGACCGCTCTCTGAACCGCAAAAGGCGCGATTC 2025937
Qy 721 CCGGCAACACATTCAGACAGCGCTGATTCGCGCTTGAACAAATTCAGCAGGAAAG 780
Db 2025936 CCGGCAACACATTCAGACAGCGCTGATTCGCGCTTGAACAAATTCAGCAGGAAAG 2025877
Qy 781 GAAAAACGCGGCAAGGCGGCAAGCTTATGTCCTTTCCAAATTA 828
Db 2025876 GAAAAACGCGGCAAGGCGGCAAGCTTATGTCCTTTCCAAATTA 2025829

RESULT 9
US-10-096-129-1
; Sequence 1, Application US/10096129
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, KARL F.
; APPLICANT: ROTH, STEPHEN
; APPLICANT: BUCZALA, STEPHANIE L.
; TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO SACCHARIDE UNITS WITH A
; FILE REFERENCE: POLYGLYCOSYLTRANSFERASE
; CURRENT APPLICATION NUMBER: US/10/096,129
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/338,943
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: US 08/478,140
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 1
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-096-129-1

Query Match      87.7%; Score 726; DB 40; Length 5859;
Best Local Similarity 93.2%; Pred. No. 6,2e-245;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 1 ATGCAAAACAGCTTATAGCTTCCGCGCAGAAAGGCGCATATGCGCAT 60
Db 1491 ATGCAAAACAGCTTATAGCTTCCGCGCAGAAAGGCGCATATGCGCAT 1550
Qy 61 ACCCTGGGAGGAGGAGGATCCGTTCACTTTTGGACGATGATCCGCTGGAAG 120
Db 1551 ACCCTGGGAGGAGGAGGATCCGTTCACTTTTGGACGATGATCCGCTGGAAG 1610
Qy 121 CTGGAACAGGCAATGGCGGAATCTGTCGCCGCTTGTGCGGCAACCCCTATTGAGCGGA 180
Db 1611 CTGGAACAGGCAATGGCGGAATCTGTCGCCGCTTGTGCGGCAACCCCTATTGAGCGGA 1670
Qy 181 GTGAAAAAACCTGCTTTATGAGCAGCCGATTTGTGAAGCAGGATGAGCAAGGT 240
Db 1671 GTGAAAAAACCTGCTTTATGAGCAGCCGATTTGTGAAGCAGGATGAGCAAGGT 1730
Qy 241 CTGCGGATATATCAACGATTTTGAAGAGAGGTTTATCGGGCAAGGTGAGAAATTC 300
Db 1731 CTGCGGATATATCAACGATTTTGAAGAGAGGTTTATCGGGCAAGGTGAGAAATTC 1790
Qy 301 CTTCGGAAGAGCCTTGCTGCAAGACGCTTTGACCCGATACCGCTTTATCTGCGC 360
Db 1791 CTTCGGAAGAGCCTTGCTGCAAGACGCTTTGACCCGATACCGCTTTATCTGCGC 1850
Qy 361 TTGGAACAGATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGAGATTACTGCGG 420
Db 1851 TTGGAACAGATGTTTATGACAGTCTGACCTCGCCCTCGGCGTGGAGATTACTGCGG 1910
Qy 421 CGGCGCTTTCGCTGTTTGAAGAGCACTGGGGGAGCGCGGCTTATCATTTCCCGA 480
Db 1911 CGGCGCTTTCGCTGTTTGAAGAGCACTGGGGGAGCGCGGCTTATCATTTCCCGA 1970
Qy 481 AAGGAGATGCGTCTTCTGCAAGGTTGCGCGCTGCGCGCGCAAGGCTGCAACCC 540
Db 1971 AAGGAGATGCGTCTTCTGCAAGGTTGCGCGCTGCGCGCGCAAGGCTGCAACCC 2030
Qy 541 GTGATCTGATGATGTTTCAAGCAATTTTTCAGACAGGAGAAATGCCGTTTCCAGCTC 600
Db 2031 GTGATCTGATGATGTTTCAAGCAATTTTTCAGACAGGAGAAATGCCGTTTCCAGCTC 2090
Qy 601 AATCCGCGCTTGTGCGCGCAAGAGCTGATATGCAAGTTTACAGCAAAACAGCGCA 660
Db 2091 AATCCGCGCTTGTGCGCGCAAGAGCTGATATGCAAGTTTACAGCAAAACAGCGCA 2150
Qy 661 TTGGGAGGCTGATTCAGACAGACCGCTCTCTGAACCGCAAAAGGCGCGATTC 720
Db 2151 TTGGGAGGCTGATTCAGACAGACCGCTCTCTGAACCGCAAAAGGCGCGATTC 2210
Qy 721 CCGGCAACACATTCAGACAGCGCTGATTCGCGCTTGAACAAATTCAGCAGGAAAG 780
Db 2211 CCGGCAACACATTCAGACAGCGCTGATTCGCGCTTGAACAAATTCAGCAGGAAAG 2270
Qy 781 GAAAAACGCGGCAAGGCGGCAAGCTTATGTCCTTTCCAAATTA 814
Db 2271 GAAAAACGCGGCAAGGCGGCAAGCTTATGTCCTTTCCAAATTA 2304

RESULT 10
US-10-654-528-1
; Sequence 1, Application US/10654528
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
```



```

; TITLE OF INVENTION: GENES ENCODING THEM
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 1
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(381)
; OTHER INFORMATION: glyls (glycyl tRNA synthetase beta chain)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (445)..(1491)
; OTHER INFORMATION: 19cA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2342)..(3262)
; OTHER INFORMATION: 19cC
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3322)..(4335)
; OTHER INFORMATION: 19cD
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4354)..(5196)
; OTHER INFORMATION: 19cE
; US-10-654-528-1

Query Match      87.7%; Score 726; DB 54; Length 5859;
Best Local Similarity 93.2%; Pred. No. 6.2e-245;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGCAAAACCAAGTATGAGCTTACGCTTCCGCGCAGAAAGGCGGCAATTCGCGAT 60
DB 1491 ATGCAAAACCAAGTATGAGCTTACGCTTCCGCGCAGAAAGGCGGCAATTCGCGAT 1550
QY 61 ACCCTCGAGGAGCGGCAATCCCGTTTTCAGTTTTCGAGCACTGATGCGCTGAAGG 120
DB 1551 ACCCTCGAGGAGCGGCAATCCCGTTTTCAGTTTTCGAGCACTGATGCGCTGAAGG 1610
QY 121 CTGGAACAGGCAATGCGGCACTCGTCCCGGCTTTCGCGCACTTATTTGAGCGGA 180
DB 1611 CTGGAACAGGCAATGCGGCACTCGTCCCGGCTTTCGCGCACTTATTTGAGCGGA 1670
QY 181 GTGGAAGAAAGCTGCTTATGAGCAAGCGGATTTGTGGAAGAGGCAATTTGAAGAGT 240
DB 1671 GTGGAAGAAAGCTGCTTATGAGCAAGCGGATTTGTGGAAGAGGCAATTTGAAGAGT 1730
QY 241 CTGCGGTATATGACCGGTATTTGAGAGCAAGTTTATCTCGGCGAAGGTGAGAAATTC 300
DB 1731 GTAACCGTATATGACCGGTATTTGAGAGCAAGTTTATCTCGGCGAAGGTGAGAAATTC 1790
QY 301 CTGCGCAAGAGAGCTTGGCTGCAAGAAAGCTTTGACCGGATATCCGCTTTATCGCGCG 360
DB 1791 CTGCGCAAGAGAGCTTGGCTGCAAGAAAGCTTTGACCGGATATCCGCTTTATCGCGCG 1850
QY 361 TTGGAAGAGATGTTATGACAGCTCTGACCTCGCTTCGCGCGGCGGAGATTAATTCGCGG 420
DB 1851 TTGGAAGAGATGTTATGACAGCTCTGACCTCGCTTCGCGCGGCGGAGATTAATTCGCGG 1910

```

```

QY 421 CGCGCTTTCCGCTGTGTTGAAAGCGAACAATGCGGAGACGCGGCTATATCATTTCCCGA 480
DB 1911 CGCGCTTTCCGCTGTGTTGAAAGCGAACAATGCGGAGACGCGGCTATATCATTTCCCGA 1970
QY 481 AAGCGATGCGGATTTTCTGGAAGGTTTCCGCGCTGCGCGCGGCGGAGGCTGACCCG 540
DB 1971 AAGCGATGCGGATTTTCTGGAAGGTTTCCGCGCTGCGCGCGGCGGAGGCTGACCCG 2030
QY 541 GTCGATCTGATGATGTTTCAAGGATTTTTCGACAGGAGAAAGATCCGTTTCCAGCTC 600
DB 2031 GTCGATCTGATGATGTTTTCGACAGGATTTTTCGACAGGAGAAAGATCCGTTTCCAGCTC 2090
QY 601 AATCCGCTTGTGCGCCCAAGAGCTGATATGACCAAGTTTCAACCAAAACAGCCCA 660
DB 2091 AATCCGCTTGTGCGCCCAAGAGCTGATATGACCAAGTTTCAACCAAAACAGCCCA 2150
QY 661 TTGGGACGCTGATGAAACAGACCGCTCTGAAACCGCAAAACAGCAAGCGGATTC 720
DB 2151 TTGGGACGCTGATGAAACAGACCGCTCTGAAACCGCAAAACAGCAAGCGGATTC 2210
QY 721 CCGGCAACATTTAAACACCGCTGATCCGCGCTTGAACCAAAATGAGAGGAGAAAG 780
DB 2211 CCGGCAACATTTAAACACCGCTGATCCGCGCTTGAACCAAAATGAGAGGAGAAAG 2270
QY 781 GAAAAACCGCGCAAGGCGGGAACGTTTCAATTG 814
DB 2271 GAAAAACCGCGCAAGGCGGGAACGTTTCAATTG 2304

RESULT 11
US-10-654-528-7
; Sequence 7, Application US/10654528
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 7
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1491)..(2330)
; OTHER INFORMATION: 19cB
; US-10-654-528-7

Query Match      87.7%; Score 726; DB 54; Length 5859;
Best Local Similarity 93.2%; Pred. No. 6.2e-245;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGCAAAACCAAGTATGAGCTTACGCTTCCGCGCAGAAAGGCGGCAATTCGCGAT 60
DB 1491 ATGCAAAACCAAGTATGAGCTTACGCTTCCGCGCAGAAAGGCGGCAATTCGCGAT 1550
QY 61 ACCCTCGAGGAGCGGCAATCCCGTTTTCAGTTTTCGAGCACTGATGCGCTGAAGG 120
DB 1551 ACCCTCGAGGAGCGGCAATCCCGTTTTCAGTTTTCGAGCACTGATGCGCTGAAGG 1610

```



```

QY 121 CTGAAACAGGCAATGGCGAAGCTGTCCTCCGCGCTTGTGCGGACCCCTATTGAGCGA 180
DB 1611 CTGGAACGGGCAATGGCGAAGCTGTCCTCCGCGCTTGTGCGGACCCCTATTGAGCGA 1670
QY 181 GTGGAAGAAAGCCTGCTTATGAGCAAGCGGATATGTGAGAGCAAGGATGGAAGT 240
DB 1671 GTGGAAGAAAGCCTGCTTATGAGCAAGCGGATATGTGAGAGCAAGGATGGAAGG 1730
QY 241 CTGCGGATATGACCGTATTGAGAGCAAGTTTACTCGGCGAAGGTGAGAAAATTC 300
DB 1731 GTACGATATGTCGCGGATTTGAGAGAGTGTACTCGCGCAAGCGCGAGAGTTCC 1790
QY 301 CTGCGGAGAGACGCTGGCTGCAAGAGCGTTGACCGCGATACCGCTTTATGTCGCG 360
DB 1791 CTGCGGAGAGATCTTGTGCTGCAAGAGCGTTGACCGCGATACCGCTTTATGTCGCG 1850
QY 361 TTGGAAGAGATGTTATGACAGTCTGACCTCGCCCTCGGCGTGGCGAATTAATGCGGG 420
DB 1851 TTGGAAGAGATGTTATGACAGTCTGACCTCGCCCTCGGCGTGGCGAATTAATGCGGG 1910
QY 421 CGGCGCTTTCGCTGTTGGAAGAGCACTGCGGAGCGCGGCTATATCATTTCCGA 480
DB 1911 CGGCGCTTTCGCTGTTGGAAGAGCACTGCGGAGCGCGGCTATATTAATTTCCGA 1970
QY 481 AAGGCAATGCGGTTTTCCTGCAAGGTTTCCGCGCTGCGCGCGCAAGGCTGCAAGCC 540
DB 1971 AAGGCAATGCGGTTTTCCTGCAAGGTTTCCGCGCTGCGCGCGCAAGGCTGCAAGCC 2030
QY 541 GTGATCTGATGATGTTGACGCAATTTTTCGACAGGAGAAATGCGGTTTTCGAGCTC 600
DB 2031 GTGATTTGATGATGTTTTCGCAAGCCTGACAGAGAGAAAGATGCGGTTTTCGAGCTC 2090
QY 601 AATCCGCTTGTGCGCGCAAGAGCTGATATGCAAGTTTCAAGCAAAACAGCGCA 660
DB 2091 AATCCGCTTGTGCGCGCAAGAGCTGATATGCAAGTTTCAAGCAAAACAGCGCA 2150
QY 661 TTGGGAGAGCTGATTCGAGACAGACCGGCTCTGACCGGCAAGCAAGAGCGGATTC 720
DB 2151 TTGGGAGAGCTGATTCGAGACAGACCGGCTCTGACCGGCAAGCAAGAGCGGATTC 2210
QY 721 CCGGCAACATTCATTAACACCGGCTGATCCGCGCTTTCGACCAAAATCAGCAGGAAAG 780
DB 2211 CCGGCAACATTCATTAACACCGGCTGATCCGCGCTTTCGACCAAAATCAGCAGGAAAG 2270
QY 781 GAAAAAGCGCGGCAAGAGCGGCAAGCAATTCATTG 814
DB 2271 GAAAAAGCGCGGCAAGAGCGGCAAGCAATTCATTG 2304

RESULT 12
US-11-102-497-1
; Sequence 1, Application US/1102497
; GENERAL INFORMATION:
; APPLICANT: Gofeschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/11/102,497
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13

```

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(381)
; OTHER INFORMATION: glys [glycy] tRNA synthetase beta chain)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (445)..(1491)
; OTHER INFORMATION: lgta
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2342)..(3262)
; OTHER INFORMATION: lgta
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3322)..(4335)
; OTHER INFORMATION: lgtd
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4354)..(5196)
; OTHER INFORMATION: lgce
US-11-102-497-1

Query Match      87.7%; Score 726; DB 70; Length 5859;
Best Local Similarity 93.2%; Pred. No. 6, 2e-245;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGCAAAACAGCTTATGAGCTTAAGTTCCGCGCAAGAGGCGGACATTTGCCGAT 60
DB 1491 ATGCAAAACAGCTTATGAGCTTGTGCGCGCAAGAGGCGGACATTTGCCGCA 1550
QY 61 ACCTTGCGAGGACGCGCATCCCGTTTCAAGTTTTCGACAGCATGATGCGTGAAGG 120
DB 1551 ACCTTGCGAGGATGCGGCAATCCCGTTTTCGACAGCATGATGCGTGAAGG 1610
QY 121 CTGGAACAGGCAATGGCGAAGTCTGTCGCGCTTGTGCGGCAAGCCTTATTTGAGCGGA 180
DB 1611 CTGGAACAGGCAATGGCGAAGTCTGTCGCGCTTGTGCGGCAAGCCTTATTTGAGCGGA 1670
QY 181 GTGGAAGAAAGCCTGCTTATGAGCAAGCGGCTATTTGTGGAAGAGGATTTGAGAGG 240
DB 1671 GTGGAAGAAAGCCTGCTTATGAGCAAGCGGCTATTTGTGGAAGAGGATTTGAGAGG 1730
QY 241 CTGCGGATATGACCGTATTGAGAGCAAGTTTACTCGGCGAAGGTGAGAAAATTC 300
DB 1731 GTACGATATGTCGCGGATTTGAGAGAGTGTACTCGCGCAAGCGCGAGAGTTCC 1790
QY 301 CTGCGGAGAGACGCTGGCTGCAAGAGCGTTGACCGCGATACCGCTTTATGTCGCG 360
DB 1791 CTGCGGAGAGATCTTGTGCTGCAAGAGCGTTGACCGCGATACCGCTTTATGTCGCG 1850
QY 361 TTGGAAGAGATGTTATGACAGTCTGACCTCGCCCTCGGCGTGGCGAATTAATGCGGG 420
DB 1851 TTGGAAGAGATGTTATGACAGTCTGACCTCGCCCTCGGCGTGGCGAATTAATGCGGG 1910
QY 421 CGGCGCTTTCGCTGTTGGAAGAGCACTGCGGAGCGCGGCTATATCATTTCCGA 480
DB 1911 CGGCGCTTTCGCTGTTGGAAGAGCACTGCGGAGCGCGGCTATATTAATTTCCGA 1970
QY 481 AAGGCAATGCGGTTTTCCTGCAAGGTTTTCGCGCTCGCGCGCAAGGCGGCTGCAAGCC 540
DB 1971 AAGGCAATGCGGTTTTCCTGCAAGGTTTTCGCGCTCGCGCGCAAGGCGGCTGCAAGCC 2030
QY 541 GTGATCTGATGATGTTGACGCAATTTTTCGACAGGAGAAATGCGGTTTTCGAGCTC 600
DB 2031 GTGATTTGATGATGTTTTCGCAAGCCTGACAGAGGAGAAATGCGGTTTTCGAGCTC 2090
QY 601 AATCCGCTTGTGCGCGCAAGAGCTGATATGCAAGTTTCAAGCAAAACAGCGCA 660

```

Db 2091 AATCCGCTTGTGGCCCAAGAGTGATTAATGCAAGTTTCAGACCAAAACAGCGCA 2150
Qy 661 TTGGGACGCTGATGGAACAGACCGCTCTTGAACCGCAACAGCAAGGCGCATTC 720
Db 2151 TTGGGACGCTGATGGAACATGACCGCTTGAACCGCAACAGCAATGGCGCATTC 2210
Qy 721 CCCGCAACACATTCGAAACACCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG 780
Db 2211 CCCGCAACACATTCGAAACACCGCTGATCCGCGCTTGACCAAAATGCGAGGAAAG 2270
Qy 781 GAAAAACGCGGCAAGGCGGCAAGTTCATTC 814
Db 2271 GAAAAACGCGGCAAGGCGGCAAGTTCATTC 2304

RESULT 13
US-11-102-497-7
; Sequence 7, Application US/11102497
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: 040853-01-5029-03
; CURRENT APPLICATION NUMBER: US/11/102,497
; PRIOR FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: US/10/654,528
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 7
; LENGTH: 5859
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1491)..(2330)
; OTHER INFORMATION: 1GCB
US-11-102-497-7

Query Match 87.7%; Score 726; DB 70; Length 5859;
Best Local Similarity 93.2%; Pred. No. 6, 2e-245;
Matches 759; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
Qy 1 ATGCAAAACCAAGTATGAGCTTAGCTTCGCGCAGAACGAGGCGGCAATTCGCGAT 60
Db 1491 ATGCAAAACCAAGTATGAGCTTAGCTTCGCGCAGAACGAGGCGGCAATTCGCGCA 1550
Qy 61 ACCTTCGAGGACGAGGATCCCGTTGAGTTTTCGAGCACTGATGCGCTGAAAAG 120
Db 1551 ACCTTCGAGGATGCGGACATCCCGTTGAGTTTTCGAGCACTGATGCGCTGAAAAG 1610
Qy 121 CTGGAACAGGCAATGCGGAACTCGTCCCGCTTGTCGCGCACCCCTATTGAGCGGA 180
Db 1611 CTGGAACAGGCAATGCGGAACTCGTCCCGCTTGTCGCGCACCCCTATTGAGCGGA 1670
Qy 181 GTGGAAAAAGCTGCTTTATGAGCCACGCGCTATTGTGGAAGCAGGCAATTCGAAAGT 240
Db 1671 GTGGAAAAAGCTGCTTTATGAGCCACGCGCTATTGTGGAAGCAGGCAATTCGAAAGT 1730
Qy 241 CTGCGCTATATGACCGTATTTGAGAGCAGCTTTTATCTGCGGAGAGTGAGAAAAATTC 300
Db 1731 GTACCGTATATGCGCGGTATTGGAAGATGATGCTTATCTGCGGAGAGGCGGAGCAGTTTC 1790

Qy 301 CTTCGCGAAGACGCTTGGCTGCAAGAACGTTTGAACCGGATACCGCTTTATGCTCCGC 360
Db 1791 CTTCGCGAAGATCTTGGCTGCAAGAACGCTTTGAACCGGATACCGCTTTATGCTCCGC 1850
Qy 361 TTGAAACGATGTTTATGACAGCTCTGACCTGCGCTTCGCGGCTGAGCAATTCGCGG 420
Db 1851 TTGAAACGATGTTTATGACAGCTCTGACCTGCGCTTCGCGGCTGAGCAATTCGCGG 1910
Qy 421 CGCGCTTTCGCTGTTGAAAGGCAACATGAGGAGGAGCGGAGCTATATATTCCTCCGA 480
Db 1911 CGCGCTTTCGCTTTCGAAAGGCAACATGAGGAGGAGCGGAGCTATATATTCCTCCGA 1970
Qy 481 AAAGGATGCGGTTTTCCTGCAAGGTTTCCGCGCTGCGCGGAAAGGCTGACCC 540
Db 1971 AAAGGATGCGGTTTTCCTGCAAGGTTTCCGCGCTGCGCGGAAAGGCTGACCC 2030
Qy 541 GTGCATCTGATGATGTTTCAGGATTTTTCGACAGGAGGAAATGCGGTTTTCAGGCTC 600
Db 2031 GTGCATCTGATGATGTTTCAGGATTTTTCGACAGGAGGAAATGCGGTTTTCAGGCTC 2090
Qy 601 AATCCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTCAGACCAAAACAGCGCA 660
Db 2091 AATCCGCTTGTGCGCCCAAGAGCTGATTAATGCAAGTTTCAGACCAAAACAGCGCA 2150
Qy 661 TTGGGACGCTGATGGAACAGACCGCTCTGGAACCGCAACAGCAAGGCGCATTC 720
Db 2151 TTGGGACGCTGATGGAACATGACCGCTCTGGAACCGCAACAGCAATGGCGCATTC 2210
Qy 721 CCCGCAACACATTCGAAACACCGCTGATCCGCGCTTGACCAAAATCAGCAGGAAAG 780
Db 2211 CCCGCAACACATTCGAAACACCGCTGATCCGCGCTTGACCAAAATGCGAGGAAAG 2270
Qy 781 GAAAAACGCGGCAAGGCGGCAAGTTCATTC 814
Db 2271 GAAAAACGCGGCAAGGCGGCAAGTTCATTC 2304

RESULT 14
US-10-467-657-1309
; Sequence 1309, Application US/10467657
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWm199, version 1.04
; SEQ ID NO 1309
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1309

Query Match 87.1%; Score 721.2; DB 51; Length 837;
Best Local Similarity 92.9%; Pred. No. 6e-244;
Matches 756; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
Qy 1 ATGCAAAACCAAGTATGAGCTTAGCTTCGCGCAGAACGAGGCGGCAATTCGCGAT 60
Db 1 ATGCAAAACCAAGTATGAGCTTAGCTTCGCGCAGAACGAGGCGGCAATTCGCGCA 60
Qy 61 ACCTTCGAGGACGAGGATCCCGTTGAGTTTTCGAGCACTGATGCGCTGAAAAG 120
Db 61 ACCTTCGAGGATGCGGAGATCCCGTTGAGTTTTCGAGCACTGATGCGCTGAAAAG 120
Qy 121 CTGGAACAGGCAATGCGGAACTCGTCCCGCTTGTCGAGGCAACCCCTATTGAGCGGA 180

```
Db 121 CTGAAACAGCGATGCGGAACTGTCCTCCGCTGTCGCGCACTTATTTGAGCGGA 180
Qy 181 GTGAAAAAAGCCCTTTATGAGCAAGCCGATTTGGAAGAGGATTTGAGCAAGG 240
Db 181 GTGAAAAAAGCCCTTTATGAGCAAGCCGATTTGGAAGAGGATTTGAGCAAGG 240
Qy 241 CTGCGGATATCAACCGTATTTGAGAGCAAGCTTTTACTCGCGAAGGCGGAAATTC 300
Db 241 CTGCGGATATCAACCGTATTTGAGAGCAAGCTTTTACTCGCGAAGGCGGAAATTC 300
Qy 301 CTGCGGAAAGACGCTTGCTGCAAGACGCTTTGACCCGATACCGCTTTATCTGCG 360
Db 301 CTGCGGAAAGATCTTGCTGCAAGACGCTTTGATAGGATTCGCTTTATCTGCG 360
Qy 361 TTGGAACCAATGTTTATGACAGCTCTGCAAGCTTCTCGCGCTGCGGATTAATGCG 420
Db 361 TTGGAACCAATGTTTATGACAGCTCTGCAAGCTTCTCGCGCTGCGGATTAATGCG 420
Qy 421 CGGCGCTTTCGCTGTTGGAAGCAAGCACTGCGGAGGCGGCTATATCATTTCCGA 480
Db 421 CGGCGCTTTCGCTGTTGGAAGCAAGCACTGCGGAGGCGGCTATATTTCCGA 480
Qy 481 AAGCGATGCGGTTTCTGCAAGGTTTCCGCTCGCGCTCGCAAGGCTGCAAGCC 540
Db 481 AAGCGATGCGGTTTCTTGAAGGTTTCCGCTCGCGCTCGCAAGGCTGCAAGCC 540
Qy 541 GTGATCTGATGATTTTCAAGCAATTTTTCAGAGGAAATGCGGTTTCCAGCTC 600
Db 541 GTGATCTGATGATTTTCAAGCAATTTTTCAGAGGAAATGCGGTTTCCAGCTC 600
Qy 601 AATCCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTCAAGCAAGCAAGCGCA 660
Db 601 AATCCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTCAAGCAAGCAAGCGCA 660
Qy 661 TTGGGAGCTGATTCAGACGACGCTTCTGCAAGCAAGCAAGCAAGGCGGATTC 720
Db 661 TTGGGAGCTGATTCAGACGACGCTTCTGCAAGCAAGCAAGCAAGGCGGATTC 720
Qy 721 CCGGCAACATTCATTAACAGCCGCTGATTCGCGCTTGAACAAATTCAGAGGAAAG 780
Db 721 CCGGCAACATTCATTAACAGCCGCTGATTCGCGCTTGAACAAATTCAGAGGAAAG 780
Qy 781 GAAAAACGCGGCAAGGCGGCAAGCTTATG 814
Db 781 GAAAAACGCGGCAAGGCGGCAAGCAATG 814
```

RESULT 15

```
US-10-467-657A-1309
/ Sequence 1309, Application US/10467657A
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON SPA
/ APPLICANT: FONTANA Maria Rita
/ APPLICANT: PIZZA Mariagrazia
/ APPLICANT: MASIGNANI Vega
/ APPLICANT: MONACI Elisabetta
/ TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE: 002441.00061
/ CURRENT APPLICATION NUMBER: US/10/467,657A
/ PRIOR FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: GB-0103424.8
/ NUMBER OF SEQ ID NOS: 9218
/ SOFTWARE: SeqMin99, version 1.04
/ SEQ ID NO 1309
/ TYPE: DNA
/ LENGTH: 837
/ ORGANISM: Neisseria gonorrhoeae
US-10-467-657A-1309
```

```
Query Match 87.1%; Score 721.2; DB 51; Length 837;
Best Local Similarity 92.9%; Pred. No. 6e-244;
```

```
Matches 756; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
Qy 1 ATGCAAAACCAAGTTATGAGCTTACGCTTCCGCGCAAGACGAGGCGCATTTGCCAT 60
Db 1 ATGCAAAACCAAGTTATGAGCTTACGCTTCCGCGCAAGACGAGGCGCATTTGCCAT 60
Qy 61 ACCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 61 ACCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Qy 121 CTGGAACAGGCAATGCGGAACTGCTCCCGCTTGTCCGCGCAAGCCCTTATTTGAGCG 180
Db 121 CTGGAACAGGCAATGCGGAACTGCTCCCGCTTGTCCGCGCAAGCCCTTATTTGAGCG 180
Qy 181 GTGAAAAAAGCCCTTTATGAGCAAGCCGATTTGGAAGAGGATTTGAGCAAGG 240
Db 181 GTGAAAAAAGCCCTTTATGAGCAAGCCGATTTGGAAGAGGATTTGAGCAAGG 240
Qy 241 CTGCGGATATCAACCGTATTTGAGAGCAAGCTTTTACTCGCGAAGGCGGAAATTC 300
Db 241 CTGCGGATATCAACCGTATTTGAGAGCAAGCTTTTACTCGCGAAGGCGGAAATTC 300
Qy 301 CTGCGGAAAGACGCTTGCTGCAAGACGCTTTGACCCGATACCGCTTTATCTGCG 360
Db 301 CTGCGGAAAGATCTTGCTGCAAGACGCTTTGATAGGATTCGCTTTATCTGCG 360
Qy 361 TTGGAACCAATGTTTATGACAGCTCTGCAAGCTTCTCGCGCTGCGGATTAATGCG 420
Db 361 TTGGAACCAATGTTTATGACAGCTCTGCAAGCTTCTCGCGCTGCGGATTAATGCG 420
Qy 421 CGGCGCTTTCGCTGTTGGAAGCAAGCACTGCGGAGGCGGCTATATCATTTCCGA 480
Db 421 CGGCGCTTTCGCTGTTGGAAGCAAGCACTGCGGAGGCGGCTATATTTCCGA 480
Qy 481 AAGCGATGCGGTTTCTGCAAGGTTTCCGCTCGCGCTCGCAAGGCTGCAAGCC 540
Db 481 AAGCGATGCGGTTTCTTGAAGGTTTCCGCTCGCGCTCGCAAGGCTGCAAGCC 540
Qy 541 GTGATCTGATGATTTTCAAGCAATTTTTCAGAGGAAATGCGGTTTCCAGCTC 600
Db 541 GTGATCTGATGATTTTCAAGCAATTTTTCAGAGGAAATGCGGTTTCCAGCTC 600
Qy 601 AATCCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTCAAGCAAGCAAGCGCA 660
Db 601 AATCCGCTTGTGCGCCCAAGAGCTGATATGCAAGTTTCAAGCAAGCAAGCGCA 660
Qy 661 TTGGGAGCTGATTCAGACGACGCTTCTGCAAGCAAGCAAGCAAGGCGGATTC 720
Db 661 TTGGGAGCTGATTCAGACGACGCTTCTGCAAGCAAGCAAGCAAGGCGGATTC 720
Qy 721 CCGGCAACATTCATTAACAGCCGCTGATTCGCGCTTGAACAAATTCAGAGGAAAG 780
Db 721 CCGGCAACATTCATTAACAGCCGCTGATTCGCGCTTGAACAAATTCAGAGGAAAG 780
Qy 781 GAAAAACGCGGCAAGGCGGCAAGCTTATG 814
Db 781 GAAAAACGCGGCAAGGCGGCAAGCAATG 814
```

Search completed: April 7, 2006, 21:06:22
Job time : 9854.41 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 17:01:35 ; Search time 3557.5 Seconds
(without alignments)
706.685 Million cell updates/sec

Title: US-09-211-691-1

Perfect score: 828
Sequence: 1 atgcaaacacgtatcacg.....tcattgtccttcacataa 828

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 6718506 seqs, 1518135245 residues

Total number of hits satisfying chosen parameters: 13437012

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents NA New: *
1: /SID55/ptodata/1/pna/PCT_NEW_COMB.seq:*
2: /SID55/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /SID55/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /SID55/ptodata/1/pna/US09_NEW_COMB.seq:*
5: /SID55/ptodata/1/pna/US10_NEW_COMB.seq:*
6: /SID55/ptodata/1/pna/US10_NEW_COMB.seq:*
7: /SID55/ptodata/1/pna/US10_NEW_COMB.seq:*
8: /SID55/ptodata/1/pna/US11_NEW_COMB.seq:*
9: /SID55/ptodata/1/pna/US11_NEW_COMB.seq:*
10: /SID55/ptodata/1/pna/US11_NEW_COMB.seq:*
11: /SID55/ptodata/1/pna/US11_NEW_COMB.seq:*
12: /SID55/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description
C 1	826.4	99.8	2242716	7 US-10-018-470B-1	Sequence 1, Appl1
C 2	36.4	4.4	2242716	7 US-10-018-470B-1	Sequence 1, Appl1
C 3	36.4	4.3	270	8 US-11-329-388-18070	Sequence 18070, A
C 4	36.4	4.3	286	10 US-11-266-748A-12825	Sequence 12825, A
C 5	36.4	4.3	286	10 US-11-266-748A-64567	Sequence 64567, A
C 6	36.4	4.3	286	10 US-11-266-748A-67339	Sequence 67339, A
C 7	36.4	4.3	1000	10 US-11-266-748A-11615	Sequence 11615, A
C 8	36.4	4.3	1000	10 US-11-266-748A-158329	Sequence 158329, A
C 9	36.4	4.3	1000	10 US-11-266-748A-286336	Sequence 286336, A
C 10	36.4	4.3	1000	10 US-11-266-748A-337765	Sequence 337765, A
C 11	36.4	4.3	1000	10 US-11-266-748A-396766	Sequence 396766, A
C 12	36.4	4.3	1000	10 US-11-266-748A-467812	Sequence 467812, A
C 13	36.4	4.3	2821	8 US-11-056-355B-89151	Sequence 89151, A
C 14	36.4	4.3	2821	8 US-11-056-355B-92907	Sequence 92907, A
C 15	35.6	4.3	516	8 US-11-330-364-14690	Sequence 14690, A
C 16	35.6	4.2	1241	10 US-11-266-748A-357498	Sequence 357498, A
C 17	35.6	4.2	1241	10 US-11-266-748A-440877	Sequence 440877, A
C 18	35.6	4.2	1704	7 US-10-531-147-2511	Sequence 2511, Ap
C 19	35.6	4.2	1704	7 US-10-531-147-2511	Sequence 3648, Ap
C 20	35.6	4.2	3593	12 US-60-742-219-2223	Sequence 2223, Ap
C 21	34.8	4.2	5157	7 US-10-461-673-7787	Sequence 7787, Ap

22	34.8	4.2	5174	10 US-11-266-748A-23500	Sequence 23500, A
23	34.2	4.1	449	8 US-11-239-610A-28539	Sequence 28539, A
24	34.2	4.1	5518	8 US-11-360-355-5183	Sequence 5183, Ap
25	34.2	4.1	1077	7 US-10-501-442-39	Sequence 39, Appl
26	34	4.1	4485	7 US-10-567-867-1977	Sequence 1977, Ap
27	34	4.1	4485	7 US-60-751-455-2720	Sequence 2720, Ap
28	33.8	4.1	5572	8 US-11-360-355-755	Sequence 755, App
29	33.8	4.1	118664	7 US-10-540-898-826	Sequence 826, App
30	33.6	4.1	7177	8 US-11-350-336-7	Sequence 7, Appl1
31	33.6	4.1	7364	8 US-11-350-336-5	Sequence 5, Appl1
32	33.6	4.1	7364	12 US-60-751-455-1947	Sequence 1947, Ap
33	33.6	4.1	7376	8 US-11-350-336-3	Sequence 3, Appl1
34	33.4	4.0	950	10 US-11-266-748A-357343	Sequence 357343, A
35	33.4	4.0	950	10 US-11-266-748A-440722	Sequence 440722, A
36	33.4	4.0	1062	7 US-10-560-224-40	Sequence 40, Appl
37	33.4	4.0	1062	7 US-10-560-414-40	Sequence 40, Appl
38	33.4	4.0	2110	6 US-10-533-515-1874	Sequence 1874, Ap
39	33.4	4.0	5834	7 US-10-567-867-2773	Sequence 2773, Ap
40	33.4	4.0	6132	7 US-10-567-867-2774	Sequence 2774, Ap
41	33.2	4.0	438	8 US-11-353-150-39324	Sequence 39324, A
42	33.2	4.0	221454	1 PCT-US05-18584-1	Sequence 1, Appl1
43	33	4.0	393	8 US-11-353-150-35925	Sequence 35925, A
44	33	4.0	576	8 US-11-329-388-9222	Sequence 9222, Ap
45	33	4.0	938	8 US-11-056-355B-1081	Sequence 1081, Ap

ALIGNMENTS

RESULT 1
US-10-018-470B-1/c
Sequence 1, Application US/10018470B
GENERAL INFORMATION:
APPLICANT: FRASER, Claire Marie
APPLICANT: HICKER, Erin Kathleen
APPLICANT: PETERSON, Jeremy D.
APPLICANT: TETTELIN, Hevry
APPLICANT: VENTER, J. Craig
APPLICANT: MASTENANT, Vega
APPLICANT: GALBOTTI, Cesira
APPLICANT: MORA, Marirosa
APPLICANT: RATTI, Giulio
APPLICANT: SCARSELLI, Maria
APPLICANT: SCARIELLO, Vincenzo
APPLICANT: RAPPUOLI, Rino
APPLICANT: PIZZA, Mariagrazia
APPLICANT: GRANDI, Guido
TITLE OF INVENTION: NEISSERIA GENOMIC SEQUENCES AND METHODS OF THEIR USE
FILE REFERENCE: 223002100400
CURRENT APPLICATION NUMBER: US/10/018,470B
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: PCT/US00/05928
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: US 60/132,068
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: GB 0004695.3
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: PCT/US99/23573
PRIOR FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 115
SOFTWARE: Seqwin99, version 1.0.4
SEQ ID NO 1
LENGTH: 2242716
TYPE: DNA
ORGANISM: Neisseria meningitidis
US-10-018-470B-1
Query Match 99.8%; Score 826.4; DB 7; Length 2242716;
Best Local Similarity 99.9%; Pred. No. 5.4e-24;
Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 ATGCAAAACGAGTATGAGCTTCCGCGAGAGCGAGGCGCACATTCGCAT 60
|||||

```

Db 1997047 ATGCAAAACAGCTTATGAGTTAGCTTCCGCCGAGAAACGAGGCGGACATTCGCCAT 1996988
QY 61 ACCCTTGGGAGGAGCGGATCCCGTTTCAGTTTTCAGAGCACTGATGCGTCTGAAAG 120
Db 1996987 ACCCTTGGGAGGAGCGGATCCCGTTTCAGTTTTCAGAGCACTGATGCGTCTGAAAG 1996928
QY 121 CTGGAACAGGAGCAATGAGCGGAACTCGTCCCGGCTTGTGGGCGGACCCCTATTTGAGCGGA 180
Db 1996927 CTGGAACAGGAGCAATGAGCGGAACTCGTCCCGGCTTGTGGGCGGACCCCTATTTGAGCGGA 1996868
QY 181 GTGAAAAAGCTGCTTTATGAGCAAGCGGATTTGTGAGCAAGGATTTGAGCAAGGAT 240
Db 1996867 GTGAAAAAGCTGCTTTATGAGCAAGCGGATTTGTGAGCAAGGATTTGAGCAAGGAT 1996808
QY 241 CTGCGGTATATCAGCGTATTTGAGGAGCAAGGATTTTACTGCGGAGAGGATGAGAAAAATTC 300
Db 1996807 CTGCGGTATATCAGCGTATTTGAGGAGCAAGGATTTTACTGCGGAGAGGATGAGAAAAATTC 1996748
QY 301 CTGCGGAAAGAGCGCTTGGCTGCAAGAACGCTTTGACCGGATACCGCTTTATTCGTCCGC 360
Db 1996747 CTGCGGAAAGAGCGCTTGGCTGCAAGAACGCTTTGACCGGATACCGCTTTATTCGTCCGC 1996688
QY 361 TTGGAACAGATTTTATGAGCAAGCTTGCCTGAGCTTGCCTCGGCGGAGGATTTACTGCGGG 420
Db 1996687 TTGGAACAGATTTTATGAGCAAGCTTGCCTGAGCTTGCCTCGGCGGAGGATTTACTGCGGG 1996628
QY 421 CGCGGCTTTCCGCTGTTGGAAGAGCAACCTGCGGAGAGCGGCGCTATATCATTTCCGCA 480
Db 1996627 CGCGGCTTTCCGCTGTTGGAAGAGCAACCTGCGGAGAGCGGCGCTATATCATTTCCGCA 1996568
QY 481 AAGAGGATGCGGTTTTCTTCTGAGCAGGTTTGCCTGCGGCTGCGCGGAGAGGCTGCAACCC 540
Db 1996567 AAGAGGATGCGGTTTTCTTCTGAGCAGGTTTGCCTGCGGCTGCGCGGAGAGGCTGCAACCC 1996508
QY 541 GTGCAATCGATGATGTTTCAAGGATTTTTCGAGAGGAGAGATGCGGTTTTGCGAGCTC 600
Db 1996507 GTGCAATCGATGATGTTTCAAGGATTTTTCGAGAGGAGAGATGCGGTTTTGCGAGCTC 1996448
QY 601 AATCCCGCCTTGTGCGGCCCAAGAGCTGATTAATGCCAAGTTTTCAGCAACCAAAACAGCGCA 660
Db 1996447 AATCCCGCCTTGTGCGGCCCAAGAGCTGATTAATGCCAAGTTTTCAGCAACCAAAACAGCGCA 1996388
QY 661 TTGGGAGGCTGATGGAACAGACGCGCTTGAACCGGAAACAGAGAGGCGGCAATTC 720
Db 1996387 TTGGGAGGCTGATGGAACAGACGCGCTTGAACCGGAAACAGAGAGGCGGCAATTC 1996328
QY 721 CCGGCGAAGCATTTCAAAACAGCGGCTGATCGCGGCTTGAACCAAAATCGAGAGGAAAG 780
Db 1996327 CCGGCGAAGCATTTCAAAACAGCGGCTGATCGCGGCTTGAACCAAAATCGAGAGGAAAG 1996268
QY 781 GAAAAACGCGGCAAAAGGCGGCAAGCTGATTTGCTTCCATTA 828
Db 1996267 GAAAAACGCGGCAAAAGGCGGCAAGCTGATTTGCTTCCATTA 1996220

```

RESULT 2
US-10-018-470B-1
; Sequence 1, Application US/10018470B
; GENERAL INFORMATION:

```

; APPLICANT: FRASER, Claire Marie
; APPLICANT: HICKEY, Erin Kathleen
; APPLICANT: PETERSON, Jeremy D.
; APPLICANT: TETTELIN, Heirve
; APPLICANT: VENTER, J. Craig
; APPLICANT: MASIGNANI, Vega
; APPLICANT: GALBOTTI, Cesira
; APPLICANT: MORA, Marirosa
; APPLICANT: RATTI, Giulio
; APPLICANT: SCARABELLI, Maria
; APPLICANT: SCARABATO, Vincenzo
; APPLICANT: RAPPUOLI, Rino
; APPLICANT: PIZZA, Mariagrazia
; APPLICANT: GRANDI, Guido

```

```

; TITLE OF INVENTION: NEISSERIA GENOMIC SEQUENCES AND METHODS OF THEIR USE
; FILE REFERENCE: 22302100400
; CURRENT APPLICATION NUMBER: US/10/018, 470B
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: PCT/US00/05928
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: US 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: GB 0004695.3
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/US99/23573
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: SeqWin99, version 1.0.4
; SEQ ID NO 1
; LENGTH: 2242716
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; US-10-018-470B-1

```

Query Match 4.4%; Score 36.4; DB 7; Length 2242716;
Best Local Similarity 51.2%; Pred. No. 19;
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

```

QY 551 TGATGTTGAGGATTTTTCAGAGGAGGAAATGCGGTTTCCAGCTCAATCCGCGCT 610
Db 920217 TGCTTACGGGAGAGTCTTCAACCGCAATTCCTCCGACCGCACTGCTGCTCCCA 920276
QY 611 TGTGCGCCCAAGAGCTGATTAATGCCAAGTTTTCAGCAACCAAAACAGGCAATTTGGCAGCC 670
Db 920277 TCAGGCGGCAAGCGCTTGAACAGCCTTGCCTGCTGATGATGCGGCGGCGGCAATTC 920336
QY 671 TGATGAAACAGACCGGCTTCTGAAACGCAACAGCAAAAGGCGCA 716
Db 920337 AAATGGGAGAGAGCTTTCTCTGCGCTGAAAGAGCTGCGCTGCA 920382

```

```

RESULT 3
US-11-329-388-18070
; Sequence 18070, Application US/11329388
; GENERAL INFORMATION:
; APPLICANT: Fisher, Scott E.
; APPLICANT: Andersen, Scott E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 16517.357 - 38-21(15503)C/US
; CURRENT APPLICATION NUMBER: US/11/329,388
; CURRENT FILING DATE: 2006-01-11
; PRIOR APPLICATION NUMBER: 09/553,094
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18831
; SEQ ID NO 18070
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB84-023-Q1-E1-A5
; US-11-329-388-18070

```

Query Match 4.3%; Score 36; DB 8; Length 270;
Best Local Similarity 49.5%; Pred. No. 0.62;
Matches 93; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```

QY 279 CCGCGAAGGTGAGAAAAATTCCTTCCGAAAGACGCTTGGCTGCAAGACGTTTGACCC 338
Db 71 CGACGAGAGCAGGTGGTATCTCTGCGAACGACCGCACTCCAGAGTGGTGAGAGT 130
QY 339 GGATACCGGCTTATCGTCCGCTTGAAGACATGTTTATGACAGCTTGCAGCCGCTC 398
Db 131 CCTGACCTGATCAGCTGCGGAGGCTGAGAGTGAAGTGGAGCTTCCCTGCTGCTGT 190
QY 399 CGGCGTGGCGATTAATGCGGCGGCGCTTCCGCTGTTGAAACGCAACTGCGGAGC 458

```

```

Db      191 CAGCAGCTCGGGCACCAGACGCTGGGGCTTTGCACGTTGGGATGAAGGGCGCTACTCGAGAC 250
QY      459 GCGGGGCT 466
          |||
          251 AGCAGGAT 258

RESULT 4
US-11-266-748A-12825/c
; Sequence 12825, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12825
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-12825

Query Match      4.3%; Score 36; DB 10; Length 286;
Best Local Similarity 49.0%; Pred. No. 0.64;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0.

QY      580 GGAATGCGGCTTGGCAGCTCAATCCCGCTTGTGCGCCCAAGAGCTGATTATGCCAAG 639
Db      247 GGAGGGGTGGCTACTTACTCTGACCAAGCTCTTGTGCGAGCTGGCCCTGCTGAGTGGCTG 188
          |||
QY      640 TTTTCAGACACAAACAGCGCATTTGGGACAGCTGATCGAACACAGCCGCTCTGAAACGGC 699
Db      187 GGTCAAGGCCACACCTCTCCACTGAGCCCGTGGCAGAGCTAGCGGCGTCTCTTCAAGCTTC 128
          |||
QY      700 AAACAGCAAGCGCGCATTTCCCGCCCAACATTTCAACACCGGCTGATCCGCGCTTGG 759
Db      127 TGGAGGACAGCGCGCTGCTGCTGCCACCCACTGCTTTCAGACACTCTCTCCAGACTACTAT 68
          |||
QY      760 ACCAAATCCAGCAGG 775
Db      67 CAGATCCACAGCATG 52

RESULT 5
US-11-266-748A-64567
; Sequence 64567, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12825
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-12825

```

```

: FILE REFERENCE: 55815-0102 (319189)
: CURRENT APPLICATION NUMBER: US/11/266, 748A
: CURRENT FILING DATE: 2005-11-03
: PRIOR APPLICATION NUMBER: EP 04105479.2
: PRIOR FILING DATE: 2004-11-03
: PRIOR APPLICATION NUMBER: EP 04105482.6
: PRIOR FILING DATE: 2004-11-03
: PRIOR APPLICATION NUMBER: EP 04105483.4
: PRIOR FILING DATE: 2004-11-03
: PRIOR APPLICATION NUMBER: EP 04105507.0
: PRIOR FILING DATE: 2004-11-03
: PRIOR APPLICATION NUMBER: EP 04105485.9
: PRIOR FILING DATE: 2004-11-03
: PRIOR APPLICATION NUMBER: EP 60/700,233
: PRIOR FILING DATE: 2005-07-18
: NUMBER OF SEQ ID NOS: 483996
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 64567
: LENGTH: 286
: TYPE: DNA
: ORGANISM: Homo Sapiens
US-11-266-748A-64567

Query Match          4.3%; Score 36; DB 10; Length 286;
Best Local Similarity 49.0%; Pred. No. 0.64;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGGTTTGGCCAGCTCAATCCGCCCTTGTGGCCCCAAGAGCTGATTAATGCAAG 639
DB 40 GGAGAGGGTGGCTACTACTGACCGACGAGCTCTCTCCAGCTGGCGCTTGAGTGGCCTG 99
QY 640 TTTCACGACCAAAACACGCGATTGGGCGACCTGATCGAACGACGCCCTCTCTGAACCGC 699
DB 100 GGTGAGGCCCAACACCTCTCCCACTGAGCCCGGTGAGAGGCTACGAGCGCTCCTCAGCCTC 159
QY 700 AAACAGCAAAAGCGCGCATTTCCCGCGCAACACATTTCAAACACGCGCTGATCCGCGCCTTG 759
DB 160 TGGAGGACGCGCCCGCTGCTGCGCACCACTGCTTCAGACACTCTCTCCAGATGAGCTTAT 219
QY 760 ACCAAATCCAGCAGGG 775
DB 220 CAGATCCGACGAGTG 235

RESULT 6
US-11-266-748A-67399/c
: Sequence 67399, Application US/11266748A
: GENERAL INFORMATION:
: APPLICANT: Hartlin, Paul
: APPLICANT: Johnston, Patrick
: TITLE OF INVENTION: Transcriptome Microarray Technology and
: FILE REFERENCE: 55815-0102 (319189)
: CURRENT APPLICATION NUMBER: US/11/266, 748A
: CURRENT FILING DATE: 2005-11-03
: PRIOR APPLICATION NUMBER: EP 04105479.2
: PRIOR FILING DATE: 2004-11-03
: PRIOR APPLICATION NUMBER: EP 04105482.6
: PRIOR FILING DATE: 2004-11-03
: PRIOR APPLICATION NUMBER: EP 04105483.4
: PRIOR FILING DATE: 2004-11-03
: PRIOR APPLICATION NUMBER: EP 04105507.0
: PRIOR FILING DATE: 2004-11-03
: PRIOR APPLICATION NUMBER: EP 04105485.9
: PRIOR FILING DATE: 2004-11-03
: PRIOR APPLICATION NUMBER: EP 04105484.2
: PRIOR FILING DATE: 2004-11-03
: PRIOR APPLICATION NUMBER: US 60/662, 746
: PRIOR FILING DATE: 2004-11-03
: PRIOR APPLICATION NUMBER: US 60/662, 746

```

/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO: 67399
/ LENGTH: 286
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-67399

Query Match 4.3%; Score 36; DB 10; Length 286;
Best Local Similarity 49.0%; Pred. No. 0.64;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGTTTGCAGCTCAATCCGCTTGTGTGCGCCAAAGAGCTGATTATGCCAAG 639
DB 247 GGAGAGGGTGTCTACTACTGACAGCCCTCTGTCCAGCTTGCCCTGTGAGTGCCCTG 188
QY 640 TTTCACGACCAAAACAGCGCATTTGGGCAAGCTGATGAAACAGACCGCTCTGAACGCG 699
DB 187 GGTCAAGGCCCAACCTCTCCACTGAGCCCCGTGAGAGCTACGGCGCTCCCTCAGCCTTC 128
QY 700 AAACAGCAAGCGCGGATTTCCCGCCCAACATTCAAACACCGCTGATCCGCGCTTG 759
DB 127 TGGGAGCAGCGCGCTGCTGCTCCACCTCTTCAGACACTCTCTCCGATGAGCTTAT 68
QY 760 ACCAAATCAGCAGG 775
DB 67 CAGGATCCCAAGTG 52

RESULT 7
US-11-266-748A-116165
/ Sequence 116165, Application US/11266748A
/ GENERAL INFORMATION:
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT APPLICATION NUMBER: US/11/266,748A
/ PRIOR FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO: 116165
/ LENGTH: 1000
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-116165

Query Match 4.3%; Score 36; DB 10; Length 1000;
Best Local Similarity 49.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 580 GGAATGCCGTTTGCAGCTCAATCCGCTTGTGTGCGCCAAAGAGCTGATTATGCCAAG 639

DB 146 GGAGAGGGTGTCTACTACTAACCAGCTCTCTGTCCAGCTTGCCCTTCTAGTGCCCTG 205
QY 640 TTTCACGACCAAAACAGCGCATTTGGGCAAGCTGATGAAACAGACCGCTCTGAACGCG 699
DB 206 GGTCAAGGCCCAACCTCTCCACTGAGCCCCGTGAGAGCTACGAGCGCTCCCTCAGCCTTC 265
QY 700 AAACAGCAAGCGCGGATTTCCCGCCCAACATTCAAACACCGCTGATCCGCGCTTG 759
DB 266 TGGGAGCAGCGCGCTGCTGCTCCACCTGCTTCAGACACTCTCTCCGATGAGCTTAT 325
QY 760 ACCAAATCAGCAGG 775
DB 326 CAGGATCCCAAGTG 341

RESULT 8
US-11-266-748A-158329/c
/ Sequence 158329, Application US/11266748A
/ GENERAL INFORMATION:
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT APPLICATION NUMBER: US/11/266,748A
/ PRIOR FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO: 158329
/ LENGTH: 1000
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-158329

Query Match 4.3%; Score 36; DB 10; Length 1000;
Best Local Similarity 49.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGTTTGCAGCTCAATCCGCTTGTGTGCGCCAAAGAGCTGATTATGCCAAG 639
DB 855 GGAGAGGGTGTCTACTACTGACAGCCCTCTGTCCAGCTTGCCCTGTGAGTGCCCTG 796
QY 640 TTTCACGACCAAAACAGCGCATTTGGGCAAGCTGATGAAACAGACCGCTCTGAACGCG 699
DB 795 GGTCAAGGCCCAACCTCTCCACTGAGCCCCGTGAGAGCTACGAGCGCTCCCTCAGCCTTC 736
QY 700 AAACAGCAAGCGCGGATTTCCCGCCCAACATTCAAACACCGCTGATCCGCGCTTG 759
DB 735 TGGGAGCAGCGCGCTGCTGCTCCACCTGCTTCAGACACTCTCTCCGATGAGCTTAT 676
QY 760 ACCAAATCAGCAGG 775
DB 675 CAGGATCCCAAGTG 660


```
RESULT 9
US-11-266-748A-286336
; Sequence 286336, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 286336
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-286336

Query Match      4.3%; Score 36; DB 10; Length 1000;
Best Local Similarity 49.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGTTTGCAGCTCAATCCGCTTGCGCCCAAGAGCTGATTATGCGAAG 639
    |||
DB 146 GGAAGAGGTGGCTACTACTGACGACGCTCTGCGCAGCTGCGCTGAGTGGGCTG 205
    |||
QY 640 TTTCAGACCAAAACAGCGCATTTGGGAGCGCTGATGAAACAGACCGCTCTGAAACGC 639
    |||
DB 206 GGTACAGGCCACACCTCCCACTGAGCGCCGTCAGAGCTACGCGGCTCCCTCAGGCTC 265
    |||
QY 700 AAACAGCAAAAGGCGCATTTCCCGGCAACACATTGAAACAGCGCGCTGATCGGCGCTTG 759
    |||
DB 266 TGGAGAGAGCGCGCTGCTGCTGCAACCTGCTTCCAGACCTCTCCGAGTACCTAT 325
    |||
QY 760 ACCAATATCAGCAGG 775
    |||
DB 326 CAGATCCAGCAGTG 341
    |||

RESULT 10
US-11-266-748A-337765/c
; Sequence 337765, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
```

```

; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 337765
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-337765

Query Match      4.3%; Score 36; DB 10; Length 1000;
Best Local Similarity 49.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGTTTGCAGCTCAATCCGCTTGCGCCCAAGAGCTGATTATGCGAAG 639
    |||
DB 855 GGAAGAGGTGGCTACTACTGACGACGCTCTGCGCAGCTGCGCTGAGTGGGCTG 796
    |||
QY 640 TTTCAGACCAAAACAGCGCATTTGGGAGCGCTGATGAAACAGACCGCTCTGAAACGC 639
    |||
DB 795 GGTACAGGCCACACCTCCCACTGAGCGCCGTCAGAGCTACGCGGCTCCCTCAGGCTC 736
    |||
QY 700 AAACAGCAAAAGGCGCATTTCCCGGCAACACATTGAAACAGCGCGCTGATCGGCGCTTG 759
    |||
DB 735 TGGAGAGAGCGCGCTGCTGCTGCAACCTGCTTCCAGACCTCTCCGAGTACCTAT 676
    |||
QY 760 ACCAATATCAGCAGG 775
    |||
DB 675 CAGATCCAGCAGTG 660
    |||

RESULT 11
US-11-266-748A-396766
; Sequence 396766, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 396766
; LENGTH: 1000
```

;; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-396766

Query Match 4.3%; Score 36; DB 10; Length 1000;
Best Local Similarity 49.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGCTTTCAGCTCAATCCCGCTTGTGGCCCAAGACATGATTATGCCAG 639
DB 146 GGAAGAGGTGGCTACTACTGACAGCCTCTCTCCAGCCTGGCCCTGTGAGTGGCTTG 205
QY 640 TTTCACGACCAAAAGAGCGCATTTGGCGACCTGATCGAACAAGCCGCTCTGAACCGC 699
DB 206 GGTTCAGGCCCAACACCTCCCACTGAGCCCGGTGACAGAGCTACGGGGCTCCTCAGCCTC 265
QY 700 AAACAGCAAGGCGGATTTCCCGCCCAACATTCAAACACCGCTGATCCGGCGCTTG 759
DB 266 TGGAGACAGCGCGCTGCTGCTGCACCCACGCTTTCAGACCTCTCCGAGTAGCTAT 325
QY 760 ACCAAATCAGCAGG 775
DB 326 CAGATCCACAGATG 341

RESULT 12

US-11-266-748A-467812/c
; Sequence 467812, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/562,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 467812
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-467812

Query Match 4.3%; Score 36; DB 10; Length 1000;
Best Local Similarity 49.0%; Pred. No. 1.1;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 580 GGAATGCCGCTTTCAGCTCAATCCCGCTTGTGGCCCAAGACATGATTATGCCAG 639
DB 855 GGAAGAGGTGGCTACTACTGACAGCCTCTCTCCAGCCTGGCCCTGTGAGTGGCTTG 796
QY 640 TTTCACGACCAAAAGAGCGCATTTGGCGACCTGATCGAACAAGCCGCTCTGAACCGC 699
DB 795 GGTTCAGGCCCAACACCTCCCACTGAGCCCGGTGACAGAGCTACGGGGCTCCTCAGCCTC 736

QY 700 AAACAGCAAGGCGGATTTCCCGCCCAACATTCAAACACCGCTGATCCGGCGCTTG 759
DB 735 TGGAGACAGCGCGCTGCTGCTGCACCACTGCTTTCAGACCTCTCCGAGTAGCTAT 676
QY 760 ACCAAATCAGCAGG 775
DB 675 CAGATCCACAGATG 660

RESULT 13

US-11-056-355B-89151
; Sequence 89151, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 89151
; LENGTH: 2821
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2821)
; OTHER INFORMATION: Ceres Seq. ID no. 12715871
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2821)
; OTHER INFORMATION: Ortholog of Ceres SEQ ID NO 13590691
; OTHER INFORMATION: as cited in SEQ ID NO 66290
US-11-056-355B-89151

Query Match 4.3%; Score 36; DB 8; Length 2821;
Best Local Similarity 51.2%; Pred. No. 1.6;
Matches 84; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 639 GTTTCAGCAAAACAGCGCATTTGGCGAGCTGATGAAACAGACCGGCTCTGAACCG 698
DB 1841 GTGGACAGCAACCAAGACAGACCTGTGAGCCTGATGAACATCTTCTTCCACAGGTCAA 1900
QY 699 CAACAGCAAAAGCGCATTTCCCGCCCAACATTCAAACACCGCTGATCCGGCGCTT 758
DB 1901 GTGAATCTAGATCGTTCTCGGACGAGCTAATGGAATTGTATGATCAGGAGTT 1960
QY 759 GACCAAAATCAGAGGAAAGGAAACCGCGCAAGGCGCG 802
DB 1961 GTATCTGCAAGATCTGAAGAAAGCTCAACCTGTGCAAAATGCGAG 2004

RESULT 14

US-11-056-355B-92907
; Sequence 92907, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 92907
; LENGTH: 2821
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:22:21 ; Search time 189 Seconds

(without alignments)
639.308 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446
Sequence: 1 MGNHYSLASAAERAHAD.....ISREKERQRHQFIVPRQ 275

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 244163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 244163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database :

A_Geneseq_21:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*
9: geneseqp2005s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1446	100.0	275	2	AAV22156 N. mening
2	1288	89.1	279	2	AAK91312 N. gonorr
3	1281	88.6	279	6	ABP77390 N. gonorr
4	1007	69.6	280	2	AAK91315 N. gonorr
5	1002	69.3	280	2	AAW06580 Lipo-olig
6	996	68.9	280	6	ABP77398 N. gonorr
7	933	64.5	280	6	ABU06084 N. mening
8	364	25.2	255	8	ADU04721 M. catarr
9	251	17.6	50	2	AAW89338 Neisseria
10	231	16.0	50	2	AAW89339 Neisseria
11	130.5	9.0	521	3	AAK91300 Human ORF
12	130.5	9.0	622	4	AAW93427 Human pol
13	130.5	9.0	622	4	AAW88452 Human men
14	130.5	9.0	622	5	ABP43477 Human sec
15	130.5	9.0	622	8	ADJ11021 Human pro
16	130.5	9.0	622	9	ADY63267 Human c10
17	122.5	8.5	1128	4	ABW62795 Drosophila
18	121	8.4	738	8	ADJ18760 Human sof
19	118.5	8.2	423	5	ABW89453 Human pol
20	118.5	8.2	517	3	AAK81790 Human ORF
21	118.5	8.2	517	4	AAW93202 Human pol
22	118.5	8.2	517	4	AAW93206 Human pol
23	118.5	8.2	517	8	ADJ30558 Human pro
24	118.5	8.2	517	8	ADJ30556 Human pro

25	118.5	8.2	560	8	ADQ21291 Adg21291 Human sof
26	118.5	8.2	578	4	AAK81257 Aag81257 Human AFP
27	118.5	8.2	579	2	AAI30812 Aay30812 Human sec
28	118.5	8.2	594	6	ABU11815 Abu11815 Human MDD
29	118.5	8.2	594	6	ABU11514 Abu11514 Human MDD
30	118.5	8.2	595	3	AAK12121 Aab12121 Hydropob
31	118.5	8.2	536	5	ABP42016 Abp42016 Human ova
32	117.5	8.1	617	8	ABW81071 Abw81071 Tumour-ag
33	113.5	7.8	527	8	ADH18902 Adh18902 Human cel
34	110	7.6	201	2	AAW20604 Aaw20604 H. pylori
35	110	7.6	739	4	ABK11417 Abk11417 Human sec
36	110	7.6	739	6	ABO00923 Abo000923 Polypept1
37	104	7.2	762	9	AEC08408 Aec08408 SCFV-B2-F
38	100.5	7.0	168	4	ABG32090 Abg32090 Novel hum
39	99.5	6.9	273	4	AAK60317 Aak60317 Helicobac
40	99	6.8	234	8	ADR09098 Adr09098 Human pro
41	98.5	6.8	273	2	AAK32543 Aay32543 H. pylori
42	98.5	6.8	444	5	AAU76669 Aau76669 Helicobac
43	98	6.8	481	6	ABP60054 Abp60054 H. pylori
44	95.5	6.6	273	4	ABK60321 Abk60321 Helicobac
45	93	6.4	109	6	ABK54961 Abk54961 Mouse IgG

ALIGNMENTS

RESULT 1
AAV22156
ID AAV22156 standard; protein; 275 AA.
XX
AC AAV22156;
XX
DT 08-SEP-1999 (first entry)
XX
DE N. meningitidis Beta-1,4-galactosyltransferase.
XX
KW Beta-1,4-galactosyltransferase; 1gFB; fusion protein; catalytic domain;
XX glycosyltransferase; accessory enzyme; nucleotide sugar formation;
XX saccharide donor; oligosaccharide synthesis;
XX carbohydrate structure development.
XX
OS Neisseria meningitidis.
XX
PN MO991224-A2.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-CA001180.
XX
PR 15-DEC-1997; 97US-0069443P.
XX
PR 14-DEC-1998; 98US-00211691.
XX
PA (CANNA) NAT RES COUNCIL CANADA.
XX
PI Gilbert M, Young NM, Wakarchuk WW;
XX WPI; 1999-395174/33.
XX DR N-PSDB; AAK84281.
XX
PT A new glycosyltransferase fusion protein useful in the enzymatic
XX synthesis of oligosaccharides.
XX
PS Example 2; Fig 2; 63pp; English.
XX
CC This sequence represents the Neisseria meningitidis Beta-1,4-
CC galactosyltransferase (also referred to as 1gfb). The invention relates
CC to a nucleic acid encoding a fusion protein that comprises a
CC glycosyltransferase catalytic domain and a catalytic domain from an
CC accessory enzyme that is involved in formation of a nucleotide sugar
CC which is a saccharide donor for a glycosyltransferase reaction. The
CC fusion protein is useful in the enzymatic synthesis of oligosaccharides.
CC The fusion proteins are able to catalyse more than one reaction involved
CC in the enzymatic synthesis. This is useful for the development of

CC therapeutic agents that have specific carbohydrate structures.
CC Carbohydrates are involved in recognition elements on the surface of
CC cells. The fusion protein can be used for the synthesis of both natural
CC carbohydrates and synthetic derivatives with novel properties. The fusion
CC polypeptide allows two glycosyltransferase reactions in a single vessel,
CC provides improved yields of end products. Additionally, cleanup and
CC disposal of extra solvents and by-products is reduced. The fusion protein
CC can also use directly different donor analogues and various acceptors
CC with a terminal galactose residue
SQ Sequence 275 AA;

Query Match 100.0%; Score 1446; DB 2; Length 275;
Best Local Similarity 100.0%; Pred. No. 3,8e-144;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MONHYISLASAERRAHIAIDTFRHGIPFPFDALMPSERLQAMALVPGISAHPIYLSG 60
DB 1 MONHYISLASAERRAHIAIDTFRHGIPFPFDALMPSERLQAMALVPGISAHPIYLSG 60
QY 61 VEKACFMSHAYLVMKQALDEGLPYITVFEDDVLLGEGEKEFLAEDAWLQERPDPTAFIVR 120
DB 61 VEKACFMSHAYLVMKQALDEGLPYITVFEDDVLLGEGEKEFLAEDAWLQERPDPTAFIVR 120
QY 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAAGIISRKAMRFFLDRAFPALPREGIHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAAGIISRKAMRFFLDRAFPALPREGIHP 180
QY 181 VDLMMFSDPFDEEGMPVQCLNPALCAQELIHYAKFHDONSALGSLIEHDRILNRKQQRDS 240
DB 181 VDLMMFSDPFDEEGMPVQCLNPALCAQELIHYAKFHDONSALGSLIEHDRILNRKQQRDS 240
QY 241 PANTFKHRLIRALTIKISREERKRRORRQFIIVPQ 275
DB 241 PANTFKHRLIRALTIKISREERKRRORRQFIIVPQ 275

RESULT 2
AAR91312
ID AAR91312 standard; protein; 279 AA.
XX
AC AAR91312;
DT 16-OCT-2003 (revised)
DT 09-JUL-1996 (first entry)
DE N. gonorrhoeae glycosyltransferase IgtB.
XX
KW Glycosyltransferase; lipo-oligosaccharide; lgt gene; LOS locus; vaccine.
XX
OS Neisseria gonorrhoeae; strain F62.
XX
FN WO9610086-A1.
XX
PD 04-APR-1996.
XX
PF 25-SEP-1995; 95WO-US012317.
XX
PR 26-SEP-1994; 94US-00312387.
XX
PA (UYRO) UNIV ROCKEFELLER.
XX
PI Gotachlich EC;
XX
DR WPI; 1996-200924/20.
XX
DR P-PSDB; AAT14061.
XX
PT Nucleic acids encoding glycosyl transferase(s) - used in the diagnosis of
XX infection with Neisseria and for the biosynthesis of oligosaccharide(s).
XX
PS Claim 9; Fig 2c; 81pp; English.
XX
CC 5 Glycosyltransferases (AAR91311-15) are products of the lgt locus

CC (AAT14061) of Neisseria gonorrhoeae strain F62. Glycosyltransferase IgtB
CC (AAR91312) can be obtained by expression of the lgtB coding sequence in
CC recombinant host cells. A method for adding Gal beta1-4 to GlcNAc or Glc
CC comprises contacting a reaction mixture containing activated Gal to an
CC acceptor moiety comprising a GlcNAc or Glc residue in the presence of
CC IgtB. Oligosaccharides can be produced that, when attached to non-toxic
CC lipids, are useful for Neisseria vaccine preparation. Blood group core
CC oligosaccharides, and mimics of lacto-N-neotetraose, gangliosides and
CC saccharide portions of globoglycolipids can also be produced using the
CC enzymes. (Updated on 16-Oct-2003 to standardise OS field)
SQ Sequence 279 AA;

Query Match 89.1%; Score 1288; DB 2; Length 279;
Best Local Similarity 90.0%; Pred. No. 2,1e-127;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;
QY 1 MONHYISLASAERRAHIAIDTFRHGIPFPFDALMPSERLQAMALVPGISAHPIYLSG 60
DB 1 MONHYISLASAERRAHIAIDTFRHGIPFPFDALMPSERLQAMALVPGISAHPIYLSG 60
QY 61 VEKACFMSHAYLVMKQALDEGLPYITVFEDDVLLGEGEKEFLAEDAWLQERPDPTAFIVR 120
DB 61 VEKACFMSHAYLVMKQALDEGLPYITVFEDDVLLGEGEKEFLAEDAWLQERPDPTAFIVR 120
QY 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAAGIISRKAMRFFLDRAFPALPREGIHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAAGIISRKAMRFFLDRAFPALPREGIHP 180
QY 181 VDLMMFSDPFDEEGMPVQCLNPALCAQELIHYAKFHDONSALGSLIEHDRILNRKQQRDS 240
DB 181 VDLMMFSDPFDEEGMPVQCLNPALCAQELIHYAKFHDONSALGSLIEHDRILNRKQQRDS 240
QY 241 PANTFKHRLIRALTIKISREERKRRORRQFIIVPQ 275
DB 241 PANTFKHRLIRALTIKISREERKRRORRQFIIVPQ 275

RESULT 3
ABP77390
ID ABP77390 standard; protein; 279 AA.
XX
AC ABP77390;
DT 07-MAR-2003 (first entry)
DE N. gonorrhoeae amino acid sequence SEQ ID 1310.
XX
KW Antibacterial; infection; vaccine; gene therapy.
XX
OS Neisseria gonorrhoeae.
XX
FN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB002069.
XX
PR 12-FEB-2001; 2001GB-00003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Piza M, Masignani V, Monaci E;
XX
DR WPI; 2003-058415/05.
XX
DR N-PSDB; ABZ38360.
XX
PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX medicament for treating or preventing N. gonorrhoeae infection.
XX
PS Disclosure; Page 283; 815pp; English.
XX
CC The present invention relates to proteins from Neisseria gonorrhoeae.

CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention
XX

SQ Sequence 279 AA;

Query Match 88.6%; Score 1281; DB 6; Length 279;
Best Local Similarity 90.0%; Pred. No. 1.1e-126;
Matches 251; Conservative 6; Mismatches 18; Indels 4; Gaps 1;

QY 1 MGNHVISLSAARRRAHIAITPGRHGIPPOFPDAMPSSRLBQAMALVPGLSAHPYLSG 60
DB 1 MGNHVISLSAARRRAHIAITPGSRGIPQFPDAMPSSRLBQAMALVPGLSAHPYLSG 60
QY 61 VEKACFMSHVAIVMEQALDDEGLPYITVPEDDVLLGEGEKEFTLADAWLQERFPDPTAFYR 120
DB 61 VEKACFMSHVAIVMEQALDDEGLPYITVPEDDVLLGEGEKEFTLADAWLQERFPDPTAFYR 120
QY 121 LETFMHVLITSPSGVADYCGRAFPILSEHMGTAGYIISRKAMRFLDRFPALPPGGLHP 180
DB 121 LETFMHVLITSPSGVADYCGRAFPILSEHMGTAGYIISRKAMRFLDRFPALPPGGLHP 180
QY 181 VDIAMFSDPFDRBGMVCCQNPALCAQELHYAKFHONSLSLIHEDRLNKKQQRDS 240
DB 181 VDIAMFSDPFDRBGMVCCQNPALCAQELHYAKFHONSLSLIHEDRLNKKQQRDS 240
QY 241 PANTFKRLIRALTKISREKRRORREQ----FIVPQ 275
DB 241 PANTFKRLIRALTKISREKRRORREQITGKIIVPQ 279

RESULT 4

AAR91315

AAR91315 standard; protein; 280 AA.

XX AAR91315;

DT 16-OCT-2003 (revised)
DT 09-JUL-1996 (first entry)

XX N. gonorrhoeae glycosyltransferase LgtE.

XX Glycosyltransferase; lipo-oligosaccharide; lgt gene; LOS locus; vaccine.

XX Neisseria gonorrhoeae; strain F62.

XX MO9610086-A1.

XX 04-APR-1996.

XX 25-SEP-1995; 95WO-US012317.

XX 26-SEP-1994; 94US-00312387.

XX (UYRO) UNIV ROCKEFELLER.

XX Gotschlich EC;

XX WPI; 1996-200924/20.

XX N-PSDB; AAT14061.

XX Nucleic acids encoding glycosyl transferase(s) - used in the diagnosis of

XX infection with Neisseria and for the biosynthesis of oligosaccharide(s).

XX Claim 12; Fig 2f; 81pp; English.

XX 5 Glycosyltransferases (AAR91311-15) are products of the lgt locus

XX (AAT14061) of Neisseria gonorrhoeae strain F62. Glycosyltransferase LgtE

XX (AAR91315) can be obtd. by expression of the lgtE coding sequence in

CC recombinant host cells. A method for adding Gal beta1-4 to GlcNAc or Glc
CC compounds contacting a reaction mixture contg. activated Gal to an
CC acceptor moiety comprising a GlcNAc or Glc residue in the presence of
CC LgtE. Oligosaccharides can be produced that, when attached to non-toxic
CC lipids, are useful for Neisseria vaccine prepn. Blood group core
CC oligosaccharides, and mimics of lacto-N-neotetraose, gangliosides and
CC saccharide portions of globoglycolipids can also be produced using the
CC enzymes. (Updated on 16-OCT-2003 to standardise OS field)
XX

SQ Sequence 280 AA;

Query Match 69.6%; Score 1007; DB 2; Length 280;
Best Local Similarity 74.7%; Pred. No. 1.2e-97;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MGNHVISLSAARRRAHIAITPGRHGIPPOFPDAMPSSRLBQAMALVPGLSAHPYLSG 60
DB 1 MGNHVISLSAARRRAHIAITPGSRGIPQFPDAMPSSRLBQAMALVPGLSAHPYLSG 60
QY 61 VEKACFMSHVAIVMEQALDDEGLPYITVPEDDVLLGEGEKEFTLADAWLQERFPDPTAFYR 120
DB 61 VEKACFMSHVAIVMEQALDDEGLPYITVPEDDVLLGEGEKEFTLADAWLQERFPDPTAFYR 120
QY 121 LETFMHVLITSPSGVADYCGRAFPILSEHMGTAGYIISRKAMRFLDRFPALPPGGLHP 180
DB 121 LETFMHVLITSPSGVADYCGRAFPILSEHMGTAGYIISRKAMRFLDRFPALPPGGLHP 180
QY 181 VDIAMFSDPFDRBGMVCCQNPALCAQELHYAKFHONSLSLIHEDRLNKKQQRDS 240
DB 181 VDIAMFSDPFDRBGMVCCQNPALCAQELHYAKFHONSLSLIHEDRLNKKQQRDS 240
QY 241 PANTFKRLIRALTKISREKRRORREQ 269
DB 237 RSLKVMFDLKRALGKFKRKKRMRORQ 265

RESULT 5

AAW06580

AAW06580 standard; protein; 280 AA.

XX AAW06580;

DT 17-OCT-2003 (revised)
DT 21-MAR-1997 (first entry)

XX lipo-oligosaccharide gene-encoded protein.

XX Polyglycosyltransferase; N-acetylglucosaminyl transferase;

XX N-acetylglucosaminyl transferase; lipo-oligosaccharide.

XX Neisseria gonorrhoeae; ATCC 33084.

XX MO9640971-A1.

XX 19-DEC-1996.

XX 03-JUN-1996; 96WO-US008323.

XX 07-JUN-1995; 95US-00478140.

XX (NEOS-) NEOSB TECHNOLOGIES INC.

XX Johnson KF, Roth S, Buczala SL;

XX WPI; 1997-052351/05.

XX N-PSDB; AAT49230.

XX Transfer of at least 2 saccharide units using polyglycosyltransferase -

XX isolated from N. gonorrhoeae, catalyses the addition of both GlcNAc and

XX GalNAc di:saccharide(s) units to a single galactose moiety.

XX Disclosure; Fig 2P-H; 38pp; English.

CC A lipooligosaccharide-encoding gene region (AA049230) of *Neisseria*
CC gonorrhoeae ATCC 33084 includes coding sequences for 5 proteins (AA006576
CC -80), one of which (AA006576) is a polylipoyltransferase that catalyses
CC the addition of GlcNAc and GalNAc disaccharides to a galactose moiety.
CC The function of the other 4 proteins is not stated in the specification.
CC (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 280 AA;

Query Match 69.3%; Score 1002; DB 2; Length 280;

Best Local Similarity 74.7%; Pred. No. 4e-97; Mismatches 45; Indels 4; Gaps 1;

Matches 201; Conservative 19; Mismatches 45; Indels 4; Gaps 1;

QY 1 MGNHVISLSAASERRAHADTFGRHGIPFPFDALMPSERLQAMAEVPGLSAHPYLSG 60
DB 1 MGNHVISLSAASERRAHADTFGRHGIPFPFDALMPSERLQAMAEVPGLSAHPYLSG 60
QY 61 VKACFMSHAYVMKQALDEGLPYITVFEDDVLGEGEKFLEADAMLOERPDPTAFIVR 120
DB 61 VKACFMSHAYVMKQALDEGLPYITVFEDDVLGEGEKFLEADAMLOERPDPTAFIVR 120
QY 121 LETMFMHVLTSFSGVADYCGRAFPILSESHMGCTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMHVLTSFSGVADYCGRAFPILSESHMGCTAGYIISRKAMRFFLDRFALPPEGGLHP 180
QY 121 LETMFMHVLTSFSGVADYCGRAFPILSESHMGCTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMHVLTSFSGVADYCGRAFPILSESHMGCTAGYIISRKAMRFFLDRFALPPEGGLHP 180
QY 181 VDLMEFSDFFDEBGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNLRKQORDS 240
DB 181 VDLMEFSDFFDEBGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNLRKQORDS 240
QY 181 VDLMEFSDFFDEBGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNLRKQORDS 240
DB 181 VDLMEFSDFFDEBGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNLRKQORDS 240
QY 241 PANTFKHRLIRALTKISREKRRORRQ 269
DB 241 PANTFKHRLIRALTKISREKRRORRQ 269
QY 237 RSLKVMFDLKRALGKRGREKKRMRORQ 265
DB 237 RSLKVMFDLKRALGKRGREKKRMRORQ 265

RESULT 6
ABP77398
ID ABP77398 standard; protein; 280 AA.

XX AC ABP77398;

XX DT 07-MAR-2003 (first entry)

XX DE N. gonorrhoeae amino acid sequence SEQ ID 1336.

XX KW Antibacterial; infection; vaccine; gene therapy.

XX OS *Neisseria gonorrhoeae*.

XX PN WO200279243-A2.

XX PD 10-OCT-2002.

XX PF 12-FEB-2002; 2002WO-IB002069.

XX PR 12-FEB-2001; 2001GB-00003424.

XX PA (CHIR-) CHIRON SPA.

XX PI Fontana MR, Pizsa M, Masignani V, Monaci E;

XX DR MPI; 2003-058415/05.

XX DR N-PSDB; AB238368.

PT New protein from *Neisseria gonorrhoeae*, useful for the manufacture of a
PT medicament for treating or preventing *N. gonorrhoeae* infection.

PS Disclosure; Page 284; 815pp; English.

XX The present invention relates to proteins from *Neisseria gonorrhoeae*.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing *N. gonorrhoeae*

CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP77396-ABP81046 represent nucleic acid
CC molecules of the invention

XX SQ Sequence 280 AA;

Query Match 68.9%; Score 996; DB 6; Length 280;

Best Local Similarity 74.3%; Pred. No. 1.7e-96; Mismatches 46; Indels 4; Gaps 1;

Matches 200; Conservative 19; Mismatches 46; Indels 4; Gaps 1;

QY 1 MGNHVISLSAASERRAHADTFGRHGIPFPFDALMPSERLQAMAEVPGLSAHPYLSG 60
DB 1 MGNHVISLSAASERRAHADTFGRHGIPFPFDALMPSERLQAMAEVPGLSAHPYLSG 60
QY 61 VKACFMSHAYVMKQALDEGLPYITVFEDDVLGEGEKFLEADAMLOERPDPTAFIVR 120
DB 61 VKACFMSHAYVMKQALDEGLPYITVFEDDVLGEGEKFLEADAMLOERPDPTAFIVR 120
QY 121 LETMFMHVLTSFSGVADYCGRAFPILSESHMGCTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMHVLTSFSGVADYCGRAFPILSESHMGCTAGYIISRKAMRFFLDRFALPPEGGLHP 180
QY 121 LETMFMHVLTSFSGVADYCGRAFPILSESHMGCTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMHVLTSFSGVADYCGRAFPILSESHMGCTAGYIISRKAMRFFLDRFALPPEGGLHP 180
QY 181 VDLMEFSDFFDEBGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNLRKQORDS 240
DB 181 VDLMEFSDFFDEBGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNLRKQORDS 240
QY 181 VDLMEFSDFFDEBGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNLRKQORDS 240
DB 181 VDLMEFSDFFDEBGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLNLRKQORDS 240
QY 241 PANTFKHRLIRALTKISREKRRORRQ 269
DB 241 PANTFKHRLIRALTKISREKRRORRQ 269
QY 237 RSLKVMFDLKRALGKRGREKKRMRORQ 265
DB 237 RSLKVMFDLKRALGKRGREKKRMRORQ 265

RESULT 7
ABU06084
ID ABU06084 standard; protein; 280 AA.

XX AC ABU06084;

XX DT 23-OCT-2003 (revised)

XX DT 22-JAN-2003 (first entry)

XX DE N. meningitidis vaccine antigen #61.

XX KW Vaccine; antigen; meningococcal disease; pathogenic bacteria; meningitis.

XX OS *Neisseria meningitidis* serogroup B.

XX PN WO200277648-A2.

XX PD 03-OCT-2002.

XX PF 22-MAR-2002; 2002WO-GB001399.

XX PR 22-MAR-2001; 2001GB-00007219.

XX PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.

XX PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX PI Robinson A, Gorringe AR, Hudson MJ, Bracegirdle P, West DM;

XX PI Oliver KJ, Kroll JS, Langford PR;

XX DR MPI; 2003-018958/01.

XX DR N-PSDB; ABX09919.

PT Identifying an antigen for manufacturing a vaccine against meningococcal
PT infection, comprises contacting antibodies with polypeptides, detecting
PT polypeptide-antibody complexes, and identifying bound polypeptides as
PT antigens.

PS Claim 37; Page 279-280; 310pp; English.

XX The invention relates to identifying an antigen comprising: (a) obtaining
CC antibodies against a commensal bacteria, or an extract from a commensal
CC bacteria; (b) contacting the antibodies with polypeptides obtained from

CC an expression library of either a commensal or a pathogenic bacteria; (c)
CC determining whether the polypeptides bind to antibodies; and (d) (where a
CC polypeptide binds to an antibody) identifying that polypeptide as an
CC antigen. Also included are: (1) a method of preparing a vaccine
CC composition, comprising identifying an antigen with the above method, and
CC combining the antigen with a carrier; (2) a vaccine composition obtained
CC by the above methods; (3) an antigen identified by the above methods; (4)
CC a polypeptide encoded by all or a part of a nucleic acid sequence
CC comprising the *Neisseria lactamica* DNA sequences detailed in the
CC specification; (5) a vector comprising the nucleic acid molecule; (6) a
CC method of preparing a composition for vaccination against infection by
CC pathogenic bacteria, comprising: (a) obtaining a first antigen from a
CC commensal *Neisseria*; (b) comparing the amino acid sequence of the first
CC antigen with the amino acid sequence of the second antigen from a
CC pathogenic bacteria, or comparing the sequence of a nucleic acid which
CC codes for the first antigen with the sequence of the nucleic acid that
CC codes for the second antigen; and if the first antigen is homologous to
CC the second antigen or if the nucleic acid sequence for the first antigen
CC is homologous to that of the second antigen, and (c) preparing a
CC composition for vaccination against bacterial infection comprising the
CC first antigen; and (7) an antibody that binds to the polypeptide antigen.
CC The method is useful in screening commensal and pathogenic bacteria for
CC previously unidentified vaccine antigens by identifying polypeptide
CC antigens that bind to sera raised against commensal bacterial proteins.
CC The polypeptide is useful as a vaccine antigen which may be used in the
CC manufacture of a medicament for vaccination against meningococcal
CC infection (e.g. meningitis). The present sequence represents an antigenic
CC protein from the pathogenic bacteria *N. meningitidis*. (Updated on 23-OCT-
CC 2003 to standardise OS field)

SQ Sequence 280 AA;

Query Match 64.5%; Score 933; DB 6; Length 280;
Best Local Similarity 69.5%; Pred. No. 8.2e-90;
Matches 187; Conservative 24; Mismatches 54; Indels 4; Gaps 1;

QY 1 MONHVISLSAARRAHIDTFGRHGIPOFPDAMPSERLBQAMALVYGLSAHYLSG 60
DB 1 MONHVISLSAARRAHIDTFGRHGIPOFPDAMPSERLBQAMALVYGLSAHYLSG 60
QY 61 VERACFMSHAVLWQKQIDBGLPYITVFEEDVLLGSEGEKFLADAWLQERPDPTAFIV 120
DB 61 VERACFMSHAVLWQKQIDBGLPYITVFEEDVLLGSEGEKFLADAWLQERPDPTAFIV 120
QY 121 LETMFHVLTPSPGVADYCGRAFPILSEHMGAGYIISRKAMRFLDRPALPPEGILH 180
DB 121 LETMFHVLTPSPGVADYCGRAFPILSEHMGAGYIISRKAMRFLDRPALPPEGILH 180
QY 181 VDLMFSPDFDRGEMVQQLNPALCAQELHYAKFHQNSALGSLIEHRLNKKOQRDS 240
DB 181 VDLMFSPDFDRGEMVQQLNPALCAQELHYAKFHQNSALGSLIEHRLNKKOQRDS 240
QY 241 PANTFKHLIRALTKISREERKRRREQ 269
DB 237 RSLKVMFDLKRALGKGRKKMERORQ 265

RESULT-8

ADL04721 .†
ID ADL04721 standard; protein; 255 AA.

AC ADL04721;

XX 06-MAY-2004 (first entry)

XX M. catarrhalis protein #487.

XX Moraxella catarrhalis; infection.

XX Moraxella catarrhalis.

XX US6673910-B1.

XX PN

PD 06-JAN-2004.
XX
XX 04-APR-2000; 2000US-00540236.
PF
XX 08-APR-1999; 99US-0128416P.
PR
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
PI Breton GL;
XX
XX WPI; 2004-178127/17.
DR
DR N-PSDB; ADL02801.
XX
PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT preparing a composition for diagnosing, preventing or treating infection
PT caused by Moraxella catarrhalis.
XX
XX Disclosure; SEQ ID NO 2407; 429bp; English.
XX
XX The invention relates to an isolated nucleic acid encoding an Moraxella
XX catarrhalis polypeptide. The nucleic acid is useful for preparing a
XX composition for diagnosing, preventing or treating infection caused by
XX Moraxella catarrhalis. The present sequence represents the amino acid
XX sequence of a M. catarrhalis protein.

SQ Sequence 255 AA;

Query Match 25.2%; Score 364; DB 8; Length 255;
Best Local Similarity 35.2%; Pred. No. 1.3e-29;
Matches 90; Conservative 49; Mismatches 103; Indels 14; Gaps 7;

QY 1 MONHVISLSAARRAHIDTFGRHGIPOFPDAMPSERLBQAMALVYGLSAHYLSG 60
DB 3 IONFVISTATKATRRHINCEFGKGIATFPAVPTDISKYAQLSIPINNRRLTDG 62
QY 61 VERACFMSHAVLWQKQIDBGLPYITVFEEDVLLGSEGEKFLAE-DAMLQERPDPTAFIV 119
DB 63 -EACGLSHVALWQKQIDBGLPYITVFEEDVLLGSEGEKFLAE-DAMLQERPDPTAFIV 117
QY 120 RLETFHVLTPSPGVADYCGRAFPILSEHMGAGYIISRKAMRFLDRPALPPEGILH 179
DB 118 KLET-WVERKHIIKAVTVLNHRQLCPLKTFHTGTAGYVISOQAKIILDYLTDAFEFF 176
QY 180 PYDLMFSPDFDRGEMVQQLNPALCAQELHYAKFHQNSALGSLIEHRLNKKOQRDS 239
DB 177 PIDHVLFDALISK--MSVLQVNPAMVQ----AHVSEDDTFKSLIETORKQVNOHRRR 230
QY 240 SPANTFKHLIRALTK 255
DB 231 TLAD-YGKRYYSIGK 245

RESULT 9

AAW89328
ID AAW89328 standard; peptide; 50 AA.

AC AAW89328;

XX 26-FEB-1999 (first entry)

XX *Neisseria meningitidis* IgTb C-terminal peptide.XX *Neisseria meningitidis*; IgTc; IgTb; beta-1,4-galactosyltransferase;

XX glycosyltransferase; proteolytic enzyme.

XX *Neisseria meningitidis*.

XX W09654331-A2.

XX 03-DEC-1998.

XX 26-MAY-1998; 98WO-IB000975.

XX XX

```

PR 27-MAY-1997; 97US-0047751P.
XX
PA (CANA ) NAT RES COUNCIL CANADA.
XX
PI Wakarchuk WW, Young NM;
XX
DR WPI, 1999-035177/03.
XX
PT Expressing high levels of glycosyltransferases - comprises use of either
PT host cells deficient in proteolytic enzymes or modified
PT glycosyltransferase genes deleted in a proteolytic recognition site.
PS
XX Example 1; Fig 8; 61pp; English.
XX
CC A method has been developed of expressing a glycosyltransferase in a host
CC cell. The method comprises introducing into the host cell a nucleic acid
CC encoding the glycosyltransferase and incubating the host cell under
CC conditions appropriate for expression of the glycosyltransferase, where
CC the host cell substantially lacks a protease that cleaves polypeptides
CC between two consecutive positively charged amino acid residues. The
CC glycosyltransferase can be used in in vitro production of oligosaccharide
CC structures which are potential therapeutic agents for use in the
CC manipulation of cell-cell recognition events, particularly adhesion of
CC bacteria and viruses to mammalian cells and leukocyte-endothelial cell
CC interaction through selectins in inflammation. The method provides more
CC readily recoverable active glycosyltransferases than prior art methods
CC involving mammalian glycosyltransferases. The present sequence represents
CC a C-terminal peptide from Neisseria meningitidis IgTB from the present
CC invention
XX
SQ Sequence 50 AA;

Query Match 17.6%; Score 255; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.6e-19;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 226 EHDRLNRKQQRDPSANTFKRLIRALTKISREKRRRRREQFIVPQ 275
DB 1 EHDRLNRKQQRDPSANTFKRLIRALTKISREKRRRRREQFIVPQ 50

RESULT 10
AAW89329 standard; peptide; 50 AA.
XX
AC AAW89329;
XX
DT 26-FEB-1999 (first entry)
XX
DE Neisseria meningitidis IgTB C-terminal peptide.
XX
KW Neisseria meningitidis; IgTC; IgTB; beta-1,4-galactosyltransferase;
KW glycosyltransferase; proteolytic enzyme.
XX
OS Neisseria meningitidis.
XX
PN WO9654331-A2.
XX
PD 03-DEC-1998.
XX
PF 26-MAY-1998; 98WO-IB000975.
XX
PR 27-MAY-1997; 97US-0047751P.
XX
PA (CANA ) NAT RES COUNCIL CANADA.
XX
PI Wakarchuk WW, Young NM;
XX
DR WPI, 1999-035177/03.
XX
PT Expressing high levels of glycosyltransferases - comprises use of either
PT host cells deficient in proteolytic enzymes or modified
PT glycosyltransferase genes deleted in a proteolytic recognition site.

```

```

XX
PS Example 1; Fig 8; 61pp; English.
XX
CC A method has been developed of expressing a glycosyltransferase in a host
CC cell. The method comprises introducing into the host cell a nucleic acid
CC encoding the glycosyltransferase and incubating the host cell under
CC conditions appropriate for expression of the glycosyltransferase, where
CC the host cell substantially lacks a protease that cleaves polypeptides
CC between two consecutive positively charged amino acid residues. The
CC glycosyltransferase can be used in in vitro production of oligosaccharide
CC structures which are potential therapeutic agents for use in the
CC manipulation of cell-cell recognition events, particularly adhesion of
CC bacteria and viruses to mammalian cells and leukocyte-endothelial cell
CC interaction through selectins in inflammation. The method provides more
CC readily recoverable active glycosyltransferases than prior art methods
CC involving mammalian glycosyltransferases. The present sequence represents
CC a C-terminal peptide from Neisseria meningitidis IgTB from the present
CC invention
XX
SQ Sequence 50 AA;

Query Match 16.0%; Score 231; DB 2; Length 50;
Best Local Similarity 94.0%; Pred. No. 1.6e-16;
Matches 47; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 226 EHDRLNRKQQRDPSANTFKRLIRALTKISREKRRRRREQFIVPQ 275
DB 1 EHDRLNRKQQRDPSANTFKRLIRALTKISREKRRRRREQFIVPQ 50

RESULT 11
AAB41900
XX
ID AAB41900 standard; protein; 521 AA.
XX
AC AAB41900;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1664 polypeptide sequence SEQ ID NO:3328.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
KW immunosuppressant; osteoporotic; antileukemic; immunosuppressant; cardiac;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antineoplastic;
KW antiviral; antibacterial; antifungal; antineoplastic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antineoplastic disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US008621.
XX
PR 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinketsu RA, Leach M;
XX
DR WPI, 2000-602362/57.

```


Search completed: April 7, 2006, 15:25:52
Job time : 193 secs

RESULT 15

ADL31021 standard; protein; 622 AA.

ADL31021;

20-MAY-2004 (first entry)

Human protein encoded by a full length cDNA clone SegID 3054.

human; medicine; signal transduction; glycoprotein; transcription;
oligo-capping method.

Homo sapiens.

EP1396543-A2.

10-MAR-2004.

07-JUL-2000; 2003EP-00025638.

08-JUL-1999; 99JP-00194486.

11-JAN-2000; 2000JP-0018774.

02-MAY-2000; 2000JP-00183865.

07-JUL-2000; 2000EP-00114089.

(REAS-) RES ASSOC BIOTECHNOLOGY.

Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H,
MPI; 2004-204755/20.

N-PSDB; ADL31020.

New oligonucleotide primers (830 cDNAs) useful for synthesizing full
length human cDNAs.

Example 1; SEQ ID NO 3054; 1340bp; English.

This invention relates to a novel primers useful for synthesizing full
length cDNA molecules that encode human proteins. Specifically, it refers
to secretory or membrane proteins that are potential therapeutic agents/
target molecules in the field of medicine, and in particular genes
encoding proteins that are associated with signal transduction,
glycoproteins and transcription. The present invention describes a method
for efficiently cloning a full length human cDNA from both the 5' and 3'
ends using the oligo-capping method. This polypeptide sequence is a full
length human protein of the invention.

Sequence 622 AA;

Query Match 9.0%; Score 130.5; DB 8; Length 622;

Best Local Similarity 25.1%; Pred. No. 0.00024;
Matches 58; Conservative 37; Mismatches 83; Indels 53; Gaps 12;

5 VISLSAARRAHINDTFRHGIPQGFAL-----MPSRLQAMAEIVPGLSAHRY--- 57

345 MINLRRRQDRERMLRALDAQETECRLVAVDGKANTSOVENLGIQMLPGY-RDPYHGR 403

58 -LSGVKACFMASHAVLMKQALDEGLPYITVFEDV-----LLGEGSEKFLAED 104

404 PLTKGELGFLSHYNTWKVEVDRGLQSLVFEDDLRFELFFKRLMNLMDYB----- 456

105 AMLQERFDPDTAFIVRLTFMFMVLTSPGVADYCGRAP---LIESEH-WGTAGYIIS 159

457 ---RGLDMDLIYVGRKMQVEH---PR-----KAVPRVRLVEADYSYMTLAYVIS 502

160 RKAMRFFELRPAALPEGLHVPVDTLMMFSPFDRBGMVPCOLNPALCAQELH 210

503 LOGARKLL---AAEPLSKMLPVDEFL-PVMFDGH--PVSSEYKAHPSLRNLH 547

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:26:11 ; Search time 40 Seconds
(without alignments)
661.490 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446
Sequence: 1 MGNHVLSLSAARRAHAD.....ISRRERKRQRREQFIVPQ 275

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1446	100.0	275	2 S70814	glycosyl transferase
2	1440	99.6	275	2 C81027	lacto-N-neotetraose
3	1324	91.6	279	2 A81971	lacto-N-neotetraose
4	1024.5	70.9	268	2 B81971	lacto-N-neotetraose
5	933	64.5	280	2 A81027	lacto-N-neotetraose
6	885	61.2	276	2 S70815	glycosyl transferase
7	454	31.4	266	2 S71025	lipopolysaccharide
8	410	28.4	302	1 A64077	lipopolysaccharide
9	287	20.5	282	1 F64091	lipopolysaccharide
10	202.5	14.0	268	2 AD3304	lipopolysaccharide
11	128.5	8.9	332	2 E71916	probable lipopolysaccharide
12	121	8.4	738	2 T00343	hypothetical prote
13	110.5	7.6	721	2 A83417	lipopolysaccharide
14	106.5	7.4	267	2 B64175	hypothetical prote
15	102	7.1	492	2 T20368	hypothetical prote
16	99.5	6.9	273	1 B64623	lipopolysaccharide
17	99	6.8	611	2 I50715	A2 isoform of vacu
18	98.5	6.8	444	2 F71916	probable lipopolys
19	93.5	6.5	404	2 C64597	lipopolysaccharide
20	93	6.4	133	2 PC1155	ig heavy chain pre
21	91	6.3	284	2 G72578	probable ABC trans
22	91	6.3	1624	2 C71129	probable reverse g
23	90	6.2	738	2 A71462	probable glucan br
24	90	6.2	5069	2 T17464	rifamycin polyketi
25	89.5	6.2	379	2 F75144	sarcosine oxidase,
26	88.5	6.1	273	2 E71916	probable lipopolys
27	88.5	6.1	725	2 AF3336	3-methyl-2-oxobuta
28	88	6.1	284	1 B64620	lipopolysaccharide
29	88	6.1	742	2 D86160	hypothetical prote

30	85	5.9	256	2 T00097	hypothetical prote
31	85	5.9	292	2 D71894	probable lipopolys
32	85	5.9	1189	2 I39711	cell protein - Agr
33	84	5.8	785	2 T19741	hypothetical prote
34	83.5	5.8	653	2 A81128	transcription anti
35	83.5	5.8	803	2 AD1282	leucyl-tRNA synthe
36	83.5	5.8	803	2 AH1653	leucyl-tRNA synthe
37	83	5.7	413	2 H88481	protein let-756 [i
38	83	5.7	9376	2 T14593	erythromycin synth
39	82.5	5.7	384	2 A12962	cellulose synthesi
40	82.5	5.7	389	2 E98320	hypothetical prote
41	82	5.7	380	2 AC0511	probable carnitine
42	82	5.7	896	2 S53990	phycobilisome anch
43	82	5.7	963	2 S43344	phosphoenolpyruvat
44	82	5.7	2027	2 S60123	hypothetical prote
45	82	5.7	2056	2 G88564	protein R10B11.1 [

ALIGNMENTS

```
RESULT 1
S70814
glycosyl transferase B (BC 2.4.-.-) - Neisseria meningitidis
C/Species: Neisseria meningitidis
C/Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #ext_change 09-Jul-2004
C/Accession: S70814
R/Version: M.P.; Hood, D.W.; Peak, I.R.A.; Virji, M.; Moxon, E.R.
Mol. Microbiol. 18, 729-740, 1995
A/Title: Molecular analysis of a locus for the biosynthesis and phase-variable expressi
A/Reference number: S70812; PMID:96414473; PMID:8817494
A/Accession: S70814
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-275 <JEN>
A/Cross-references: UNIPROT:Q51116; UNIPARC:UP1000016FCOF; EMBL:U25839; NID:9973183; PI
A/Genetics:
A/Keyword: glycosyltransferase
C/Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match      100.0%; Score 1446; DB 2; Length 275;
Best Local Similarity 100.0%; Pred. No. 6e-122;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MGNHVLSLSAARRAHADTFGRHGIFPQFDALMPSRLDQAAELVGLSAHPYLSG 60
      |||
DB      1 MGNHVLSLSAARRAHADTFGRHGIFPQFDALMPSRLDQAAELVGLSAHPYLSG 60

QY      61 VRKACPMASHAVVMKQALDEGLPYITVPEBDVVLGGEKEFTLADNMLQRRPDTAFYR 120
      |||
DB      61 VRKACPMASHAVVMKQALDEGLPYITVPEBDVVLGGEKEFTLADNMLQRRPDTAFYR 120

QY      121 LETMFMHVLTSPSGVADYCGRAFPILSESHMGCTAGYISRKAMRPFILDFEALPPRGGLP 180
      |||
DB      121 LETMFMHVLTSPSGVADYCGRAFPILSESHMGCTAGYISRKAMRPFILDFEALPPRGGLP 180

QY      181 VDLMPSPDFDEEGMPVQCLNPALCAQELHYAKFDONSALGSLIEHDLNRRKQQRDS 240
      |||
DB      181 VDLMPSPDFDEEGMPVQCLNPALCAQELHYAKFDONSALGSLIEHDLNRRKQQRDS 240

QY      241 PANTFKHRLIRALTISRRERKRQRREQFIVPQ 275
      |||
DB      241 PANTFKHRLIRALTISRRERKRQRREQFIVPQ 275

RESULT 2
C81027
lacto-N-neotetraose biosynthesis glycosyl transferase IgTB NMB1928 [imported] - Neisseri
C/Species: Neisseria meningitidis
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #ext_change 09-Jul-2004
C/Accession: C81027
R/Version: H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
```

Hickey, B.K.; Haft, B.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
rt, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: C81027
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-275 <TET>
A:Cross-references: UNIPROT:Q51116; UNIPARC:UPI000012E5FD; GB:AE002541; GB:AE002098; NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1928
C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 99.6%; Score 1440; DB 2; Length 275;
Best Local Similarity 99.6%; Pred. No. 2,1e-121;
Matches 274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MONHVLSASAERRAHADTEGRHGIPOFPDAMPSERLQAMALVPGLSAHPYLSG 60
DB 1 MONHVLSASAERRAHADTEGRHGIPOFPDAMPSERLQAMALVPGLSAHPYLSG 60
QY 61 VEKACFMASHAVLMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERPDPTAFYR 120
DB 61 VEKACFMASHAVLMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERPDPTAFYR 120
QY 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
QY 181 VDLMPSPDFDEGMPVQOLNPAICAOELHYAKFHONSALSLIEHDLRLNRKQORRDS 240
DB 181 VDLMPSPDFDEGMPVQOLNPAICAOELHYAKFHONSALSLIEHDLRLNRKQORRDS 240
QY 241 PANTFKRLIRALTISRERKRORRQRFIVPQ 275
DB 241 PANTFKRLIRALTISRERKRORRQRFIVPQ 275

RESULT 3
A81971
lacto-N-neotetraose biosynthesis glycosyl transferase NMA0525 [imported] - *Neisseria men*
C:Species: *Neisseria meningitidis*
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: A81971
R:Parthill, J.; Achtmann, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
J.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: A81971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-279 <PAR>
A:Cross-references: UNIPROT:P57033; UNIPARC:UPI000012E5FC; GB:AL162753; GB:AL157959; NID
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: lgtB; NMA0525
C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 91.6%; Score 1324; DB 2; Length 279;
Best Local Similarity 92.5%; Pred. No. 5,2e-111;
Matches 258; Conservative 2; Mismatches 15; Indels 4; Gaps 1;
QY 1 MONHVLSASAERRAHADTEGRHGIPOFPDAMPSERLQAMALVPGLSAHPYLSG 60
DB 1 MONHVLSASAERRAHADTEGRHGIPOFPDAMPSERLQAMALVPGLSAHPYLSG 60
QY 61 VEKACFMASHAVLMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERPDPTAFYR 120
DB 61 VEKACFMASHAVLMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERPDPTAFYR 120

QY 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
QY 181 VDLMPSPDFDEGMPVQOLNPAICAOELHYAKFHONSALSLIEHDLRLNRKQORRDS 240
DB 181 VDLMPSPDFDEGMPVQOLNPAICAOELHYAKFHONSALSLIEHDLRLNRKQORRDS 240
QY 241 PANTFKRLIRALTISRERKRORRQ---FIVPQ 275
DB 241 PANTFKRLIRALTISRERKRORRQIGKTIIVPQ 279

RESULT 4
B81971
lacto-N-neotetraose biosynthesis glycosyl transferase NMA0527 [imported] - *Neisseria men*
C:Species: *Neisseria meningitidis*
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: B81971
R:Parthill, J.; Achtmann, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
J.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: B81971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-268 <PAR>
A:Cross-references: UNIPROT:Q9JW65; UNIPARC:UPI00000C0C06; GB:AL162753; GB:AL157959; NID
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: lgtA2; NMA0527
C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 70.9%; Score 1024.5; DB 2; Length 268;
Best Local Similarity 73.6%; Pred. No. 3,3e-84;
Matches 209; Conservative 13; Mismatches 37; Indels 25; Gaps 4;
QY 1 MONHVLSASAERRAHADTEGRHGIPOFPDAMPSERLQAMALVPGLSAHPYLSG 60
DB 1 MONHVLSASAERRAHADTEGRHGIPOFPDAMPSERLQAMALVPGLSAHPYLSG 60
QY 61 VEKACFMASHAVLMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERPDPTAFYR 120
DB 61 VEKACFMASHAVLMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERPDPTAFYR 120
QY 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
QY 181 VDLMPSPDFDEGMPVQOLNPAICAOELHYAKFHONSALSLIEHDLRLNRKQORRDS 240
DB 181 VDLMPSPDFDEGMPVQOLNPAICAOELHYAKFHONSALSLIEHDLRLNRKQORRDS 240
QY 241 PANTFKRLIRALTISRERKRORRQ-----FIVPQ 275
DB 229 P-----KVLKQALGKIGRERKRORRQKLEKHLGHHVDFE 268

RESULT 5
A81027
lacto-N-neotetraose biosynthesis glycosyl transferase lgtE NMB1926 [imported] - *Neisseria men*
C:Species: *Neisseria meningitidis*
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: A81027
R:Petrelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, B.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
rt, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: A81027
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-280 <TET>
 A:Cross-references: UNIPROT:Q51117; UNIPARC:UPI00000510B1; GB:AE002541; GB:AE002098; NID
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NMB1926
 C:Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 64.5%; Score 933; DB 2; Length 280;
 Best Local Similarity 69.5%; Pred. No. 5.5e-76;
 Matches 187; Conservative 24; Mismatches 54; Indels 4; Gaps 1;

1 MONHVLSLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLRQMAELVPGLSAHNYLSG 60
 1 MONHVLSLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLRQMAELVPGLSAHNYLSG 60
 61 VERACFMSHAYLWKQALDGLPYITVPEDDVLLGSEGEKFLADAWLQERPDPTAFYR 120
 61 VERACFMSHAYLWKQALDGLPYITVPEDDVLLGKQAEKFLADWTLEERPDQSAFYR 120
 121 LETMFHVLVTPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGLHP 180
 121 LETMFHVLVTPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGLHP 180
 181 VDMMEFTYFDEKGMVQVNPALCTQELHYAKFHSKNSMLGSDLEKD---REQRRHR 236
 181 VDMMEFTYFDEKGMVQVNPALCTQELHYAKFHSKNSMLGSDLEKD---REQRRHR 236
 241 PANTFPHRLIRALTISRERKRRQRRQ 269
 241 PANTFPHRLIRALTISRERKRRQRRQ 269
 237 RSLKVMFDLKRALGKFKRKKRMRQRQ 265

RESULT 6

S70815
 glycoyl transferase B (EC 2.4.-.-) - Neisseria meningitidis
 C:Species: Neisseria meningitidis
 C>Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004
 C:Accession: S70815
 R:Jennings, M.P.; Hood, D.W.; Peak, I.R.A.; Virji, M.; Moxon, E.R.
 Mol. Microbiol. 18, 729-740, 1995
 A:Title: Molecular analysis of a locus for the biosynthesis and phase-variable expression
 A:Reference number: S70812; MID:96414473; PMID:8617494
 A:Accession: S70815
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-276 <JEN>
 A:Cross-references: UNIPROT:Q51117; UNIPARC:UPI000016FC10; EMBL:U25839; NID:G973183; PID
 C:Genetics:
 A:Gene: lgtE
 C:Superfamily: lipopolysaccharide biosynthesis-associated protein
 C:Keywords: glycosyltransferase

Query Match 61.2%; Score 885; DB 2; Length 276;
 Best Local Similarity 67.3%; Pred. No. 1.1e-71;
 Matches 181; Conservative 23; Mismatches 57; Indels 8; Gaps 2;

1 MONHVLSLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLRQMAELVPGLSAHNYLSG 60
 1 MONHVLSLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLRQMAELVPGLSAHNYLSG 60
 61 VERACFMSHAYLWKQALDGLPYITVPEDDVLLGSEGEKFLADAWLQERPDPTAFYR 120
 61 VERACFMSHAYLWKQALDGLPYITVPEDDVLLGSEGEKFLADAWLQERPDPTAFYR 120
 121 LETMFHVLVTPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGLHP 180
 121 LETMFHVLVTPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGLHP 180
 117 LETMFHVLVTPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGLHP 176
 181 VDMMEFTYFDEKGMVQVNPALCTQELHYAKFHSKNSMLGSDLEKD---REQRRHR 240
 181 VDMMEFTYFDEKGMVQVNPALCTQELHYAKFHSKNSMLGSDLEKD---REQRRHR 240

DB 177 VDMMEFTYFDEKGMVQVNPALCTQELHYAKFHSKNSMLGSDLEKD---REQRRHR 232
 QY 241 PANTFPHRLIRALTISRERKRRQRRQ 269
 DB 233 RSLKVMFDLKRALGKFKRKKRMRQRQ 261

RESULT 7

S71025
 lipopolysaccharide biosynthesis protein lfc2B - Haemophilus influenzae (strain RM7004)
 C:Species: Haemophilus influenzae
 C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S71025
 R:High, N.J.; Jennings, M.P.; Moxon, E.R.
 Mol. Microbiol. 20, 165-174, 1996
 A:Title: Tandem repeats of the tetramer 5'-CAAT-3' present in lfc2A are required for ph
 A:Reference number: S71024; MID:97014379; PMID:8861214
 A:Accession: S71025
 A:Molecule type: DNA
 A:Residues: 1-266 <HIG>
 A:Cross-references: UNIPROT:Q57394; UNIPARC:UPI000012B66A; EMBL:Z54182; NID:G1150402; PI
 A:Experimental source: strain RM7004
 C:Genetics:
 A:Gene: lfc2B
 C:Function:
 A:Description: required for the biosynthesis of a phase-variable lipopolysaccharide str

Query Match 31.4%; Score 454; DB 2; Length 266;
 Best Local Similarity 39.7%; Pred. No. 4.2e-33;
 Matches 110; Conservative 42; Mismatches 107; Indels 18; Gaps 6;

QY 3 MNVLSLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLRQMAELVPGLSAHNYLSG 62
 DB 4 NVVLSLSAABRRHIAIDTGRHGIPPOFPDAMPSSRLRQMAELVPGLSAHNYLSG 62
 QY 63 KACFMSHAYLWKQALDGLPYITVPEDDVLLGSEGEKFLADAWLQERPDPTAFYR 122
 DB 63 KACFMSHAYLWKQALDGLPYITVPEDDVLLGSEGEKFLADAWLQERPDPTAFYR 122
 QY 123 TFMFHVLTSPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGLHP 182
 DB 123 TFMFHVLTSPSGVADYCGRAFPILSESHWGAGYIISRKAMPFLDRPALPPEGLHP 182
 QY 183 LMMEBDFDRBSMPVQALNPALCAQELHYAKFHSKNSMLGSDLEKD---LNRKQQRDS 240
 DB 183 LMMEBDFDRBSMPVQALNPALCAQELHYAKFHSKNSMLGSDLEKD---LNRKQQRDS 240
 QY 183 LMMEBDFDRBSMPVQALNPALCAQELHYAKFHSKNSMLGSDLEKD---LNRKQQRDS 240
 DB 183 LMMEBDFDRBSMPVQALNPALCAQELHYAKFHSKNSMLGSDLEKD---LNRKQQRDS 240
 QY 241 PANTFPHRLIRALTISRERK--RQRRQRRQRRQ 275
 DB 233 ---TLKTVLISLAGKPKKILKIKYKLFISKATVFR 266

RESULT 8

A64077
 lipopolysaccharide biosynthesis protein lfc2A - Haemophilus influenzae
 C:Species: Haemophilus influenzae
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: A64077; S15282; S39576
 R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirnness, E.F.; Kerlavage, J.
 ; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, C.
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.; Venter,
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MID:95350630; PMID:7542800
 A:Accession: A64077
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-302 <TIGR>
 A:Cross-references: UNIPROT:Q03974; UNIPARC:UPI000012E5B0; GB:U32736; GB:L42023; NID:915
 A:Experimental source: strain Rd K120
 R:Cope, L.D.; Yoge, R.; Mertsola, J.; Latimer, J.L.; Hanson, M.S.; McCracken Jr., G.H.;

Mol. Microbiol. 5, 1113-1124, 1991
A>Title: Molecular cloning of a gene involved in lipopolysaccharide biosynthesis and virulence in *Yersinia enterocolitica*
A/Reference number: S15282; MUID:92065807; PMID:1956289
A/Accession: S15282
A/Molecule type: DNA
A/Residues: 1-28, 'N', '30-40, 45-105, 'L', '107-150, 'R', '152-302 <COP>
A/Cross-references: UNIPARC:UP1000016F7AB; EMBL:X56903; NID:948811; PIDD:CAA40221.1; PIDD:CAA40221.1; PIDD:CAA40221.1
A/Experimental source: strain D142
A/Note: It is uncertain whether Met-1 or Met-11 is the initiator
R/High, N.J.; Deadman, M.B.; Moxon, E.R.
Mol. Microbiol. 9, 1275-1282, 1993
A>Title: The role of a repetitive DNA motif (5'-CAAT-3') in the variable expression of virulence in *Yersinia enterocolitica*
A/Reference number: S39576; MUID:95020659; PMID:7523834
A/Accession: S39576
A/Molecule type: DNA
A/Residues: 1-28, 'N', '30-40, 49-105, 'L', '107-150, 'R', '152-255, 'P', '257-273, 'D', '275-302 <HIS>
A/Cross-references: UNIPARC:UP1000016F7B1; EMBL:J19441; NID:9305379; PIDD:AAA65534.1; PIDD:AAA65534.1; PIDD:AAA65534.1
A/Experimental source: strain RM7004
C/Genetics:
A/Function:
A/Description: involved in lipopolysaccharide biosynthesis and virulence expression
C/Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 28.4%; Score 410; DB 1; Length 302;
Best Local Similarity 32.2%; Pred. No. 4.3e-29;
Matches 96; Conservative 59; Mismatches 113; Indels 30; Gaps 5;
1 MONVHSLASAEARRAHADTFGRHGIPOFPDALT-----MPSERL 41
4 INIVSMENATERRRHITKOFESKLSFSPFAAYTQOSINOSINOSINOSINOSI 63
42 EOMAM--LVPLSHPTISGVKACPMASHAVLMKQALDEGLPYTTFEDVLLGEGE 98
64 NOSINSNLSINNIIESRILTKGKCLISHFLMKCNENPEYLIKIEDVLLGENE 123
99 KPLAEDATLOERPDPTATVLETFMAY-LTSPGAVADYCGRAPPLSEHWGTAGYI 157
124 VFLNOMEWKTRDFDIFIRLETFLQPVKEKQKRIPEPNSRNDDIKSTWGTAGYI 183
158 ISRKARFPLDRPALPPEGLHPVDLMFSDPFRGMPVQOLNPAALCAQELHYAKFHO 217
184 ISGAKAYVIEYKNIIPSEIYAVDELIFNKLDVQNYIYQNLPAICIQEL---QANOS 240
218 NSALGSLIEHDLNLRKQQRSDPANTFPRKRLRALTKISREKRRQRREQPIVFO 275
241 KSVLTGSLERKQKRSKIRK---KTLKQRLTRIKENIRALNKKMKQKRIKEMQ 294

RESULT 9
F64091
lipopolysaccharide biosynthesis protein lic2B homolog - Haemophilus influenzae (strain R
C/Species: Haemophilus influenzae
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C/Accession: F64091
R/Flatschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.
; Gockayne, J.D.; Scott, J.; Shirley, J.; Liu, L.; Glodok, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Pine, L.D.; Fritchman, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A./Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A/Reference number: F64091
A/Accession: F64091
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-282 <TIGR>
A/Cross-references: UNIPROT:Q57125; UNIPARC:UP10000139C94; GB:U32760; GB:L42023; NID:915
C/Genetics:
A/Start codon: GTC
C/Superfamily: lipopolysaccharide biosynthesis-associated protein

Query Match 20.5%; Score 297; DB 2; Length 282;
Best Local Similarity 31.0%; Pred. No. 5.3e-19;

Matches 86; Conservative 44; Mismatches 115; Indels 32; Gaps 11;
3 NHVHSLASAEARRAHADTFGRHGIPOFPDALTMPSERLBAMALVPGLSAH--YISG 60
34 NYVHSLTEQRRHITKOFESKLSFSPFAAYTQOSINOSINOSINOSINOSI 92
61 VEKAFMSHVAIWMKQALDEGLPYTTFEDVLLGEGEKPLAEDAMLOERPDPTATVIR 120
93 GEIGALSHIVLMQALLENNTYINIFEDDITLGENAKELLEID-----YISDDIHYLK 146
121 LETFMHVLTPSPGAVDCGR-APPLSEHWGTAGYIISRKARFPLDRPALPPEGLH 179
147 LEANGKFFKQPKSVK--CDNNVTEYVYKQ--GAGIVYTKGAKYLLIYKNEPLD--V 201
180 PVDLMFSDPFRGMPVQOLNPAALCAQELHYAKFHDONSALGSLIE-HDPLNKKQQR 238
202 AVDSIVFEDFLHFKQYKIVQLSPGICVQDF--VLHPDNPRESSIOEGRDV--HGNQRK 256
239 DSPANTFGRHLIRALTKISREKRRQRREQPIVFO 275
257 SILEKINERFGRVYKIMFGKQ-----VPRK 282

RESULT 10
AD3304
lipopolysaccharide biosynthesis protein lic2B [imported] - Brucella melitensis (strain
C/Species: Brucella melitensis
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C/Accession: AD3304
R/Delevecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Lee, T.; Ivanova,
; Mazur, M.; Goldstein, E.; Selkov, E.; Elzer, P.H.; Haglund, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A/Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A/Reference number: AD3252; PMID:11756688
A/Accession: AD3304
A/Molecule type: DNA
A/Status: preliminary
A/Residues: 1-268 <KID>
A/Cross-references: UNIPROT:O81YMS; UNIPARC:UP10000057C9D; GB:AE008917; PIDD:AAL51599.1;
A/Experimental source: strain 164
A/Genetics:
A/Map position: 1

Query Match 14.0%; Score 202.5; DB 2; Length 268;
Best Local Similarity 28.4%; Pred. No. 1.4e-10;
Matches 74; Conservative 34; Mismatches 88; Indels 65; Gaps 11;
1 MONVHSLASAEARRAHADTFGRHGIPOFPDALTMPSERLBAMALVPGLSAHFY--- 57
9 MKCYLINDKSRDRLEFMAOFRLEGAOF-----ERYAVNGRAMSPLELASFTQI 59
58 -----LSGVKACPMH-AVLMKQALDEGLPYTTFEDVLLGEGEKPLAEDAMLOE 109
60 SKEMWAPLSPALIGFLSHRCKLEKLAGED-AVAVFEHDDIRLSQSSRFPLASDHMI-- 116
110 RDPDPTATVLETFMAYLTPSPGAVDCGR-APPLSEHWGTAGYIISRKARFPL-- 167
117 ---PKQADIVKIDAGHEVLIS-NPYKMGPSISRLSRRLQOTGGYVVRDARQLPL 172
168 -----DRPALPPEGLHPVDLMFSDPFRGMPVQOLNPAALCAQELHYAKFHDONS 219
173 MEKVASPVVHFLPDPDQ-----PENNDF-----EIVQISPALCRQS-----GMS 212
220 ALGSLIEHDLNLRKQQRSDS 240
213 TIGQ-----NRRKQKPS 225

RESULT 11
E71916
probable lipopolysaccharide biosynthesis protein - *Helicobacter pylori* (strain J99)
C/Species: *Helicobacter pylori*

A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: E71916
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
I: Ives, C.; Gibson, R.; Meberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; PMID:99120557; PMID:9923682
A:Accession: E71916
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-332 <NN>
A:Cross-references: UNIPROT:Q9ZLL7; UNIPARC:UPI00000D3659; GB:AE001489; GB:AE001439; NID
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0562

Query Match 8.9%; Score 128.5; DB 2; Length 332;
Best Local Similarity 23.5%; Pred. No. 0.0081;
Matches 74; Conservative 39; Mismatches 99; Indels 103; Gaps 17;

QY 29 FQEPDLMSEK-LEQMAELVGLSA-----HPY-----LSGVKACFMSHAIVLMKOL 77
DB 35 FQFPAISPKHODFEKLELYNAOSLQSDWYHSYVAGLTLPELGCYLSHYLWKECV 94
QY 78 DEGLPYITVFEDVDVLGEGEKEFL-AEDAMLOERPDPTAFIVRL-----ETWFEMV 128
DB 95 KLDQF-VVLEDDVTL---ESHFMQALBDCLSKSPD-----FVRLGCVTYHTETKF-NV 144
QY 129 LTPSPSVA---DYCGRAPELLES-----ETWFEMV 148
DB 145 L--PKFVFPFPPDHSFKNNPILEKFKFPDVSFNLSTHKVHYILKKIQSYVATHEK 202
QY 149 -----EHM-----GTGGYIISRKAMPFLDRPALPPEGLHPVDLMSDFPBBGM 195
DB 203 EAPFLEHFLITSVYVASTAGYITLPAGAKTPIEATSP--KIIEPDMFMDSAVH---- 256
QY 196 PVCQNLPAQAEIHYAKHDONSALGSLIEHDRLNLRQQRSDSPANTFKRLIPALTK 255
DB 257 DVANLTYVCPVSL-----SHSLDSTIIQKQKSLKSYLPPOKSTIFKN-----LRY 304
QY 256 ISREKRRQRREQF 270
DB 305 YSLNAKKRLNARQY 319

RESULT 12

T00343
hypothetical protein KIAA0584 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C:Accession: T00343
R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.
DNA Res. 5, 31-39, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. IX. The complet
A:Reference number: Z14086; PMID:98290545; PMID:9628581
A:Accession: T00343
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-738 <NAG>
A:Cross-references: UNIPROT:O60327; UNIPARC:UPI00000D284; EMBL:AB011156; NID:G3043691;
A:Experimental source: brain
C:Genetics:
A:Note: KIAA0584

Query Match 8.4%; Score 121; DB 2; Length 738;
Best Local Similarity 23.1%; Pred. No. 0.01;
Matches 53; Conservative 42; Mismatches 100; Indels 34; Gaps 9;

QY 5 VLSLSAARRAHIDTGRHGIPPOFPAL-----MPSRLQMAELVVG-----LSAHP 56
DB 458 MNLKRRKDRDRMLRTTLYEOBIEVKIVAVDVGKALNTSQLALNTEMLPGRYDPYSSRP 517

QY 57 YLSGEKACFMSHAIVLMKQALDEGLPYITVFEDVDVLGEGEKEFLAE--DAWLOERFDDP 114
DB 518 LTRG-RIGCFILSHSVKMEVIDRELEKTVIEDDDRFHQFKGKMKMDNDQADWBE 576
QY 115 TAFIVALETFMNVLTSPBGVADYCGRAFPILLESB-WGTAGYIISRKAMPFLDRPAL 173
DB 577 LTVIRKRMQVQEPKAVPNVAN-----LVEADYSVYTLGYSILBGAQKLV--GAN 626
QY 174 PREGHPVDLM-----FSDFPDREGMPVQQLNALCAQELHY 211
DB 627 FPGKQLPVDEFUPVMYNHPVAEYKEYESRDLKAFSAEP-LTIYPTHY 674

RESULT 13

lpsA protein [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AH3417
R:DelVecchio, V.G.; Kapural, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella meliten
A:Reference number: AD3252; PMID:11756688
A:Accession: AH3417
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-721 <KUR>
A:Cross-references: UNIPROT:Q8Y36; UNIPARC:UPI0000058017; GB:AE008917; PIDN:AA152507.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11326
A:Map position: I

Query Match 7.6%; Score 110.5; DB 2; Length 721;
Best Local Similarity 23.2%; Pred. No. 0.087;
Matches 45; Conservative 28; Mismatches 68; Indels 53; Gaps 7;

QY 5 VLSLSAARRAHIDTGRHGIPPOF-----PDALMBSEBLEQ 43
DB 25 VINMASQPAAYKTVAASIEAYGQGFQLRIDAVNGHTATRGIDDAFDALNG----- 78
QY 44 AMAELVPGLSAHPYLSGVEKACFMSHAIVLMKQALDEGLPYITVFEDVDVLGEGEKEFLAE 103
DB 79 --REMLPG-----EYGCYRSHSKALSFSLSGSPYGLILEDDVVFETTTARLRHD 126
QY 104 DAWLOERPDPTAFIVRL-----ETWFEMVLTSPSGVADYCGRAFPILLESBHWGTAGYIIS 159
DB 127 IIKSLPDPD-----VVKLVNHRSPFLFMSLLETDA--DRIGRA---IHGQSSAAAYLVS 176
QY 160 RKAMPFLDRPAL 173
DB 177 REGARKLISALSTM 190

RESULT 14

B64175
hypothetical protein H1697 (lsg locus) - Haemophilus influenzae
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: B64175; S27580
R:Pletschmann, R.D.; Adame, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kellavagge, A
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodex, A.; Kelley, J.M.; Weidman, J
D.M.; Branden, R.C.; Fine, L.D.; Fritschman, J.L.; Fritschman, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; PMID:95350630; PMID:7542800
A:Accession: B64175
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-267 <TIG>
A:Cross-references: UNIPROT:P71398; UNIPARC:UPI00001799AC; GB:LA42023; TIGR:H1697

A:Experimental source: strain Rd KM20
R:McLaughlin, R.; Abu Kwak, Y.; Spinoia, S.; Apicella, M.
submitted to the EMBL Data Library, June 1992
A:Description: Characterization and sequence of the lsg locus from Haemophilus influenza
A:Reference number: S27577
A:Accession: S27580
A:Molecule type: DNA
A:Residues: 11-26, 'K', 28-39, 'I', 41-55, 'S', 57-111, 'L', 113-136, 'N', 138-139, 'N', 141-143, 'T'
A:Cross-references: UNIPARC:UPI000016F78F; EMBL:M94855; NID:g148931; PIDN:AAA24981.1; PI
A:Experimental source: strain A2

Query Match 7.4%; Score 106.5; DB 2; Length 267;

Best Local Similarity 23.4%; Pred. No. 0.057; Indels 43; Gaps 14;

Matches 60; Conservative 46; Mismatches 107; Indels 43; Gaps 14;

QY 1 MGNHVISLASAERRAHADTF--GRHGIPQFPDAMPSERLQAMAE--VPGISAP 56
DB 12 LKKTILSLDKDQIR---ELFFSQKNTEDFQVFSALNTQKWDLAIFNIEQFKAH- 66
QY 57 YLSGVEK--ACFMSHAVIMKQALDEG---LPYITVPEDDYLLGGEKEFLAEDAMLOE 109
DB 67 YFRNVTKGRIIGCTLSHYOKIVEDNDIABDSYALVCEDDALFHSDFQNLTL--ALLSE 124
QY 110 RRDPTDAFI-----VRLETWFMHVLTPSGVADYCGRAFPILSEHWGTAG 155
DB 125 KLBAREIILLGQSKINDPNDPDEINYPITF-SFLCKKTGDVNY--AFP-YKSYFAGTVG 179
QY 156 YIISRRAMPFLDRFAPALPPEGLHPVDLMMFSDPFRBGMFVCO---NPALCAQELHY 211
DB 180 YLIKSGAARFICQISQNKRFWLAD-DFLIFBQNFIRNKVRPLIVIEPVLIS-NLES 237
QY 212 AKFHQNSALSLIEH 227
DB 238 VRGSLSNMLKLMKY 253

RESULT 15

T20368
hypothetical protein D2045_9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T20368

R:Lloyd, C.

submitted to the EMBL Data Library, August 1994

A:Reference number: Z19262

A:Accession: T20368

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-492 <WIL>

A:Cross-references: UNIPROT:Q18989; UNIPARC:UPI0000079DA6; EMBL:Z35639; PIDN:CAA84699.1;

A:Experimental source: clone D2045

C:Genetics:

A:Gene: CESP:D2045.9

A:Map position: 3

A:Introns: 87/2; 140/3; 291/2; 367/2

Query Match

Best Local Similarity 7.1%; Score 102; DB 2; Length 492;

Matches 47; Conservative 35; Mismatches 88; Indels 52; Gaps 9;

QY 4 HVISLASAERRAHADTFGRHGIPOFPDAMPSERLQAMAE-----VPGISA 54
DB 293 YLVNLKRQERLNMOKIFDLISIEYSLER-TDQKLDLPELKNYQILLEGIDPISK 351
QY 55 HPYLSGVEKACFMSHAVIMKQALDEGLPYITVPEDDYLLGGEKEFLAEDAMLOERFDPD 114
DB 352 RPKMGK-BIGCFLSHYRWQVVOHNYEKVIVFEDDL-----RPSHD 392
QY 115 TAPYRLETMMHVLTPSGVADYCGRA-----PPLSEBHWGTAGYIIS 159
DB 393 GLTRIRVYLQDLDSHKPMDLI-YGRRKQSENEELMISQRRHLSVFEYSYV-TLGYMLS 450
QY 160 RKAMRFPLDRFAPALPPEGLHPVD--LMMFSDPFRBGMFVCO 198

DB 451 LMGARKLL--RPNPLKQWVPVDEYLPIMFNKHPNKVSLKIC 489

Search completed: April 7, 2006, 15:30:35
Job time : 42 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:22:56 ; Search time 230 Seconds

(without alignments)
843.566 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446

Sequence: 1 MGNHVISLMSAERRAHAD.....ISRREKRRGRQRTVPFQ 275

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1440	99.6	275	1 LGTB_NEIME	O5116 neisseria m
2	1427	98.7	275	2 OGR92_NEIME	O8K92 neisseria m
3	1399	96.7	279	2 O7WUX5_NEISU	O7WUX5 neisseria s
4	1392	96.3	275	2 O8L2V3_NEIME	O8L2V3 neisseria m
5	1391	96.2	275	2 O8L2V5_NEIME	O8L2V5 neisseria m
6	1386	95.9	275	2 O8L2V7_NEIME	O8L2V7 neisseria m
7	1361	94.1	275	2 O9EVD4_NEISU	O9EVD4 neisseria s
8	1360	94.1	279	2 O8K1A8_NEIME	O8K1A8 neisseria m
9	1353	93.6	279	2 O7WUX2_NEIPO	O7WUX2 neisseria p
10	1333	92.2	279	2 O8L2V9_NEIME	O8L2V9 neisseria m
11	1328	91.8	275	2 O8K98_NEIME	O8K98 neisseria m
12	1327	91.8	275	2 O8L2V2_NEIME	O8L2V2 neisseria m
13	1325	91.6	279	2 O8L2V6_NEIME	O8L2V6 neisseria m
14	1324	91.6	279	1 LGTB_NEIMA	B5703 neisseria m
15	1320	91.3	275	2 O93E5_NEIME	O93E5 neisseria m
16	1281	89.1	279	1 LGTB_NEIGO	O50947 neisseria g
17	1281	88.6	279	2 O5F4Y6_NEIG1	O5F4Y6 neisseria g
18	1273	88.0	279	2 O7WUX9_NEILA	O7WUX9 neisseria l
19	1250	86.4	279	2 O93PR9_NEIGO	O93PR9 neisseria g
20	1187	82.1	248	2 O9RGN1_NEIGO	O9RGN1 neisseria g
21	1122.5	77.6	268	2 O8K91_NEIME	O8K91 neisseria m
22	1055	73.0	280	2 O93PR7_NEIGO	O93PR7 neisseria g
23	1044.5	72.2	268	2 O8K15_NEIME	O8K15 neisseria m
24	1033.5	71.5	268	2 O8K97_SWIME	O8K97 neisseria m
25	1024.5	70.9	268	2 O7BB1_NEIME	O7BB1 neisseria m
26	1020	70.5	280	2 O9JW65_NEIMA	O9JW65 neisseria m
27	1017.5	70.4	268	2 O8L2V8_NEIME	O8L2V8 neisseria m
28	1017.5	70.4	268	2 O7WUX7_NEILA	O7WUX7 neisseria l
29	1017.5	70.4	268	2 O8L2V9_NEIME	O8L2V9 neisseria m
30	1013.5	70.1	268	2 O7WUX4_NEISU	O7WUX4 neisseria s
31	1007	69.6	280	1 LGTB_NEIGO	O50950 neisseria g

32	1003	69.4	280	2 O9REX4_NEIGO	O9REX4 neisseria g
33	1002.5	69.3	268	2 O7WUX1_NEIPO	O7WUX1 neisseria p
34	998	69.0	280	2 P96947_NEIME	P96947 neisseria m
35	996.5	68.9	268	2 O8K19_NEIME	O8K19 neisseria m
36	996	68.9	280	2 O5F4Y4_NEIG1	O5F4Y4 neisseria g
37	995.5	68.8	268	2 O8L2V8_NEIME	O8L2V8 neisseria m
38	969	67.0	280	2 O9RGM8_NEIME	O9RGM8 neisseria g
39	962	66.5	280	2 O8L2V1_NEIGO	O8L2V1 neisseria m
40	961	66.5	280	2 O8K92_NEIME	O8K92 neisseria m
41	951	65.8	280	2 O93E4_NEIME	O93E4 neisseria m
42	935	64.7	280	2 O9EVD3_NEISU	O9EVD3 neisseria s
43	933	64.5	280	1 LGTB_NEIME	O5117 neisseria m
44	932	64.5	280	2 O8L2V4_NEIME	O8L2V4 neisseria m
45	464	32.1	266	2 O8RKU0_HAEIN	O8RKU0 haemophilus

ALIGNMENTS

RESULT 1

LGTB_NEIME STANDARD; PRT; 275 AA.

AC O5116;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Lacto-N-neotetraose biosynthesis glycosyl transferase 1gtb
 (EC 2.-.-.-).
 GN Name=1gtb; OrderedLocusNames=MNB1928;
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=96414473; PubMed=8817494;
 RA Jennings M.P., Hood D., Peak I.R.A., Virji M., Moxon E.R.,
 RT "Molecular analysis of a locus for the biosynthesis and phase-variable
 expression of the lacto-N-neotetraose terminal lipopolysaccharide
 structure in Neisseria meningitidis.";
 RT Mol. Microbiol. 18:729-740(1995).
 RL [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=2015755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;
 RA Tetelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
 RA Nelson K.E., Risen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
 RA Dodson R.J., Nelson W.C., Gwinn M.L., Deboy R.T., Peterson J.D.,
 RA Hickey R.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
 RA Dougherty B.A., Mason T.M., Clecko A., Parksey D.S., Blair B.,
 RA Clifton H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
 RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Masiatani V., Pizzo M.,
 RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappunli R.,
 RA "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58.";
 RT Science 287:1809-1815(2000).
 RL [3]
 CC -1- FUNCTION: Adds the terminal galactose to the lacto-N-tetraose chain
 in lipopolysaccharide (LOS).
 CC -1- PATHWAY: Biosynthesis of the terminal lacto-N-neotetraose LPS
 structure.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase 25 family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL; U25839; AAC44085.1; -; Genomic DNA.
 CC EMBL; AE002098; AAF42257.1; -; Genomic DNA.
 CC PIR; C81027; C81027.

DR PIR: S70814; S70814.
 DR TIGR: NMB1928; -.
 DR InterPro: IPR002654; Glyco_trans_25.
 DR PANTHER: PTHR15862; Glyco_trans_25; 1.
 DR Pfam: PF01755; Glyco_trans_25; 1.
 DR Complete proteome; Glycosyltransferase;
 KM Lipopolysaccharide biosynthesis; transferase.
 FT CONFLICT 97 A -> E (in Ref. 1).
 FT MATCHES 97
 SQ SEQUENCE 275 AA; 31578 MW; E871305E2F6CF0F CRC64;

Query Match 99.6%; Score 1440; DB 1; Length 275;
 Best Local Similarity 99.6%; Pred. No. 2.2e-118;
 Matches 274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MONHVISLAAERRAHIAIDTGRGHIIPQFPDAMPSERLBOAMAEVPGLSAHPIYLSG 60
 DB 1 MONHVISLAAERRAHIAIDTGRGHIIPQFPDAMPSERLBOAMAEVPGLSAHPIYLSG 60
 QY 61 VEKACFMSHVAIVMKQALDEGLPYITVFEDDVLLGEGEKKFLAEDAWLQERFDPDTAFIVR 120
 DB 61 VEKACFMSHVAIVMKQALDEGLPYITVFEDDVLLGEGEKKFLAEDAWLQERFDPDTAFIVR 120
 QY 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGIISRKAMRFFLDRFALPPEGGLHP 180
 DB 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGIISRKAMRFFLDRFALPPEGGLHP 180
 QY 181 VDLMMFSDPFREGMPVCOINPALCAOELHYAKFHDONSALGSLIEHRLNLRKQQRDS 240
 DB 181 VDLMMFSDPFREGMPVCOINPALCAOELHYAKFHDONSALGSLIEHRLNLRKQQRDS 240
 QY 241 PANTFKHRLIRALTIXISREKRQRORRQFIIVPQ 275
 DB 241 PANTFKHRLIRALTIXISREKRQRORRQFIIVPQ 275

RESULT 2
 Q8KR92_NEIME PRELIMINARY; PRT; 275 AA.
 ID Q8KR92;
 AC Q8KR92;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Beta 1.4 galactosyltransferase.
 GN Name=lgTB;
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BB305-TT4;
 RK MEDLINE=22311028; PubMed=12423771; DOI=10.1016/S0928-8244(02)00379-6;
 RA Zhu P., Tsai C.M., Frasch C.E.;
 RT "Immunologic and genetic characterization of lipooligosaccharide
 RT variants in a Neisseria meningitidis serogroup C strain."
 RL FEMS Immunol. Med. Microbiol. 34:193-200(2002).
 DR EMBL: AY039684; F: AK85139.1; -: Genomic DNA.
 DR GO: GO:0016757; F: transferase activity; transferring glycosyl. . .; IEA.
 DR InterPro: IPR002654; P: Lipopolysaccharide biosynthesis; IEA.
 DR Pfam: PF01755; Glyco_trans_25; 1.
 DR Glycosyltransferase; Transferase.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 275 AA; 31502 MW; 9C8CB6180A719A3B CRC64;

Query Match 98.7%; Score 1427; DB 2; Length 275;
 Best Local Similarity 98.5%; Pred. No. 3e-117;
 Matches 271; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MONHVISLAAERRAHIAIDTGRGHIIPQFPDAMPSERLBOAMAEVPGLSAHPIYLSG 60
 DB 1 MONHVISLAAERRAHIAIDTGRGHIIPQFPDAMPSERLBOAMAEVPGLSAHPIYLSG 60
 QY 61 VEKACFMSHVAIVMKQALDEGLPYITVFEDDVLLGEGEKKFLAEDAWLQERFDPDTAFIVR 120

DB 61 VEKACFMSHVAIVMKQALDEGLPYITVFEDDVLLGEGEKKFLAEDAWLQERFDPDTAFIVR 120
 QY 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGIISRKAMRFFLDRFALPPEGGLHP 180
 DB 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGIISRKAMRFFLDRFALPPEGGLHP 180
 QY 181 VDLMMFSDPFREGMPVCOINPALCAOELHYAKFHDONSALGSLIEHRLNLRKQQRDS 240
 DB 181 VDLMMFSDPFREGMPVCOINPALCAOELHYAKFHDONSALGSLIEHRLNLRKQQRDS 240
 QY 241 PANTFKHRLIRALTIXISREKRQRORRQFIIVPQ 275
 DB 241 PANTFKHRLIRALTIXISREKRQRORRQFIIVPQ 275

RESULT 3
 Q7WUX5_NEISU PRELIMINARY; PRT; 279 AA.
 ID Q7WUX5;
 AC Q7WUX5;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE lgTB.
 GN Name=lgTB;
 OS Neisseria subflava.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=28449;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Zhu P., Tsai C.M.;
 RT "Galactosyltransferase gene family in Neisseria: Implication for
 RT understanding the evolution of prokaryotic enzymes responsible for
 RT glycosylation."
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY134877; AAN08515.1; -: Genomic DNA.
 SQ SEQUENCE 279 AA; 31861 MW; FDI8AAA32E155A1 CRC64;

Query Match 96.7%; Score 1399; DB 2; Length 279;
 Best Local Similarity 96.4%; Pred. No. 9.1e-115;
 Matches 269; Conservative 1; Mismatches 5; Indels 4; Gaps 1;

QY 1 MONHVISLAAERRAHIAIDTGRGHIIPQFPDAMPSERLBOAMAEVPGLSAHPIYLSG 60
 DB 1 MONHVISLAAERRAHIAIDTGRGHIIPQFPDAMPSERLBOAMAEVPGLSAHPIYLSG 60
 QY 61 VEKACFMSHVAIVMKQALDEGLPYITVFEDDVLLGEGEKKFLAEDAWLQERFDPDTAFIVR 120
 DB 61 VEKACFMSHVAIVMKQALDEGLPYITVFEDDVLLGEGEKKFLAEDAWLQERFDPDTAFIVR 120
 QY 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGIISRKAMRFFLDRFALPPEGGLHP 180
 DB 121 LETMFMHVLTPSPGVADYCGRAFPILSEHMGTAGIISRKAMRFFLDRFALPPEGGLHP 180
 QY 181 VDLMMFSDPFREGMPVCOINPALCAOELHYAKFHDONSALGSLIEHRLNLRKQQRDS 240
 DB 181 VDLMMFSDPFREGMPVCOINPALCAOELHYAKFHDONSALGSLIEHRLNLRKQQRDS 240
 QY 241 PANTFKHRLIRALTIXISREKRQRORRQFIIVPQ 275
 DB 241 PANTFKHRLIRALTIXISREKRQRORRQFIIVPQ 279

RESULT 4
 Q8L2V3_NEIME PRELIMINARY; PRT; 275 AA.
 ID Q8L2V3;
 AC Q8L2V3;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE lgTB.
 GN Name=lgTB;

```
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M992;
RX MEDLINE=22051050; PubMed=12055303;
RA Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT "Genetic diversity of three igt loci for biosynthesis of
  lipooligosaccharide (LOS) in Neisseria species.";
RL Microbiology 148:1833-1844(2002).
EMBL: AF470660; AM33865.1; -; Genomic DNA.
GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro: IPR002654; Glyco trans 25.
DR Pfam: PF01755; Glyco_transf_25; 1.
SQ SEQUENCE 275 AA; 31474 MW; 76C0BFA029DC9226 CRC64;

Query Match 96.3%; Score 1392; DB 2; Length 275;
Best Local Similarity 96.7%; Pred. No. 3.7e-114;
Matches 266; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 MONHVTSLASAEARRAHADTGRHGIPQFPDAMPSERLQAMALVPGLSAHPLYSG 60
DB 1 MONHVTSLASAEARRAHADTGRHGIPQFPDAMPSERLQAMALVPGLSAHPLYSG 60
QY 61 VERACFMSHVAVMKQALDEGLPYITVFEDVLLGEGEEKFLADAWLQRRFPDPTAFIVR 120
DB 61 VERACFMSHVAVMKQALDEGLPYITVFEDVLLGEGEEKFLADAWLQRRFPDPTAFIVR 120
QY 121 LETMFHMTVLTSSGVADYCGRAFPILSESHMGTAGYIISRKMRFFLDLRFALPPEGGLP 180
DB 121 LETMFHMTVLTSSGVADYCGRAFPILSESHMGTAGYIISRKMRFFLDLRFALPPEGGLP 180
QY 121 VDLMMFSDPFDEBGMVPCQALNPALCAQELHYAKFHDONSALGSLIEHDLRNKQOQRDS 240
DB 121 VDLMMFSDPFDEBGMVPCQALNPALCAQELHYAKFHDONSALGSLIEHDLRNKQOQRDS 240
QY 181 VDLMMFSDPFDEBGMVPCQALNPALCAQELHYAKFHDONSALGSLIEHDLRNKQOQRDS 240
DB 181 VDLMMFSDPFDEBGMVPCQALNPALCAQELHYAKFHDONSALGSLIEHDLRNKQOQRDS 240
QY 241 PANTFGRRLIRALTYSRERKRQRROQFIVPQ 275
DB 241 PANTFGRRLIRALTYSRERKRQRROQFIVPQ 275

RESULT 5
Q8L2V5 NEIME PRELIMINARY; PRT; 275 AA.
AC Q8L2V5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE LgtB.
GN Name=LgtB;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M981;
RX MEDLINE=22051050; PubMed=12055303;
RA Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT "Genetic diversity of three igt loci for biosynthesis of
  lipooligosaccharide (LOS) in Neisseria species.";
RL Microbiology 148:1833-1844(2002).
EMBL: AA470659; AM33862.1; -; Genomic DNA.
GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro: IPR002654; Glyco trans 25.
DR Pfam: PF01755; Glyco_transf_25; 1.
SQ SEQUENCE 275 AA; 31543 MW; 6EB00EB09A128DE CRC64;

Query Match 96.2%; Score 1391; DB 2; Length 275;
Best Local Similarity 96.7%; Pred. No. 4.5e-114;
Matches 266; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 1 MONHVTSLASAEARRAHADTGRHGIPQFPDAMPSERLQAMALVPGLSAHPLYSG 60
DB 1 MONHVTSLASAEARRAHADTGRHGIPQFPDAMPSERLQAMALVPGLSAHPLYSG 60
QY 61 VERACFMSHVAVMKQALDEGLPYITVFEDVLLGEGEEKFLADAWLQRRFPDPTAFIVR 120
DB 61 VERACFMSHVAVMKQALDEGLPYITVFEDVLLGEGEEKFLADAWLQRRFPDPTAFIVR 120
QY 121 LETMFHMTVLTSSGVADYCGRAFPILSESHMGTAGYIISRKMRFFLDLRFALPPEGGLP 180
DB 121 LETMFHMTVLTSSGVADYCGRAFPILSESHMGTAGYIISRKMRFFLDLRFALPPEGGLP 180
QY 121 VDLMMFSDPFDEBGMVPCQALNPALCAQELHYAKFHDONSALGSLIEHDLRNKQOQRDS 240
DB 121 VDLMMFSDPFDEBGMVPCQALNPALCAQELHYAKFHDONSALGSLIEHDLRNKQOQRDS 240
QY 181 VDLMMFSDPFDEBGMVPCQALNPALCAQELHYAKFHDONSALGSLIEHDLRNKQOQRDS 240
DB 181 VDLMMFSDPFDEBGMVPCQALNPALCAQELHYAKFHDONSALGSLIEHDLRNKQOQRDS 240
QY 241 PANTFGRRLIRALTYSRERKRQRROQFIVPQ 275
DB 241 PANTFGRRLIRALTYSRERKRQRROQFIVPQ 275
```

```
RESULT 6
Q8L2V7 NEIME PRELIMINARY; PRT; 275 AA.
AC Q8L2V7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE LgtB.
GN Name=LgtB;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=891;
RX MEDLINE=22051050; PubMed=12055303;
RA Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT "Genetic diversity of three igt loci for biosynthesis of
  lipooligosaccharide (LOS) in Neisseria species.";
RL Microbiology 148:1833-1844(2002).
EMBL: AF470658; AM33859.1; -; Genomic DNA.
GO: GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro: IPR002654; Glyco trans 25.
DR Pfam: PF01755; Glyco_transf_25; 1.
SQ SEQUENCE 275 AA; 31530 MW; 5C9BBD41BAAB04F CRC64;

Query Match 95.9%; Score 1386; DB 2; Length 275;
Best Local Similarity 96.4%; Pred. No. 1.2e-113;
Matches 265; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 1 MONHVTSLASAEARRAHADTGRHGIPQFPDAMPSERLQAMALVPGLSAHPLYSG 60
DB 1 MONHVTSLASAEARRAHADTGRHGIPQFPDAMPSERLQAMALVPGLSAHPLYSG 60
QY 61 VERACFMSHVAVMKQALDEGLPYITVFEDVLLGEGEEKFLADAWLQRRFPDPTAFIVR 120
DB 61 VERACFMSHVAVMKQALDEGLPYITVFEDVLLGEGEEKFLADAWLQRRFPDPTAFIVR 120
QY 121 LETMFHMTVLTSSGVADYCGRAFPILSESHMGTAGYIISRKMRFFLDLRFALPPEGGLP 180
DB 121 LETMFHMTVLTSSGVADYCGRAFPILSESHMGTAGYIISRKMRFFLDLRFALPPEGGLP 180
QY 121 VDLMMFSDPFDEBGMVPCQALNPALCAQELHYAKFHDONSALGSLIEHDLRNKQOQRDS 240
DB 121 VDLMMFSDPFDEBGMVPCQALNPALCAQELHYAKFHDONSALGSLIEHDLRNKQOQRDS 240
QY 181 VDLMMFSDPFDEBGMVPCQALNPALCAQELHYAKFHDONSALGSLIEHDLRNKQOQRDS 240
DB 181 VDLMMFSDPFDEBGMVPCQALNPALCAQELHYAKFHDONSALGSLIEHDLRNKQOQRDS 240
QY 241 PANTFGRRLIRALTYSRERKRQRROQFIVPQ 275
DB 241 PANTFGRRLIRALTYSRERKRQRROQFIVPQ 275
```

```

RESULT 7
Q9EVD4 NEISU
ID Q9EVD4 NEISU PRELIMINARY; PRT; 275 AA.
AC Q9EVD4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Galactosyl transferase.
GN Name=1gtB;
OS Neisseria subflava.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxId=28449;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21142520; PubMed=11208792; DOI=10.1128/JB.183.3.934-941.2001;
RT Arking D., Tong Y., Stein D.C.;
RL "Analysis of lipooligosaccharide biosynthesis in the Neisseriaceae.";
RL J. Bacteriol. 183:934-941(2001).
DR EMBL; AF240672; AAC09765.1; -; Genomic DNA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro; IPR002654; Glyco_trans_25.
DR Pfam; PF01755; Glyco_transf_25; 1.
KW Transferase.
SQ
SEQUENCE 275 AA; 31441 MW; D719F3B15F64D14C CRC64;

Query Match 94.1%; Score 1361; DB 2; Length 275;
Best Local Similarity 94.9%; Pred. No. 2e-111;
Matches 261; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 MONHVISLASAABERRAHADTFGRGIPPOFPDAMPSEERLEQMAELVPGISAHPIYLSG 60
DB 1 MONHVISLASAABERRAHADTFGRGIPPOFPDAMPSEERLEQMAELVPGISAHPIYLSG 60
QY 61 VEKACFMSHAVLMKQALDEGLPYITVFEDDVLLGGESEKFLAEDAMLOERPDPTAFIVR 120
DB 61 VEKACFMSHAVLMKQALDEGLPYITVFEDDVLLGGESEKFLAEDAMLOERPDPTAFIVR 120
QY 121 LETMFMAHVLTPSPGVADYCGRAFPFLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
DB 121 LETMFMAHVLTPSPGVADYCGRAFPFLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
QY 121 LETMFMAHVLTPSPGVADYCGRAFPFLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
DB 121 LETMFMAHVLTPSPGVADYCGRAFPFLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
QY 181 VDLMMFSDPFDREGMPVCQALNPALCAQELHYAKFHONSAAGSLIEHDLINRKQQRDS 240
DB 181 VDLMMFSDPFDREGMPVCQALNPALCAQELHYAKFHONSAAGSLIEHDLINRKQQRDS 240
QY 241 PANTFKHRLIRALTISRERERKRRORRQFIVPFQ 275
DB 241 PANTFKHRLIRALTISRERERKRRORRQFIVPFQ 275

RESULT 8
Q8K1A8 NEIME
ID Q8K1A8 NEIME PRELIMINARY; PRT; 279 AA.
AC Q8K1A8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE LgtB.
GN Name=1gtB;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxId=487;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=M120; and 7880;
RX MEDLINE=22051050; PubMed=12055303;
RT Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT "Genetic diversity of three lgt loci for biosynthesis of
RT lipooligosaccharide (LOS) in Neisseria species.";

```

```

RL Microbiology 148:1833-1844(2002).
DR EMBL; AF470662; AAM3870.1; -; Genomic DNA.
DR EMBL; AF470663; AAM3872.1; -; Genomic DNA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro; IPR002654; Glyco_trans_25.
DR Pfam; PF01755; Glyco_transf_25; 1.
SQ
SEQUENCE 279 AA; 31660 MW; 099D2E1E0DF456F6 CRC64;

Query Match 94.1%; Score 1360; DB 2; Length 279;
Best Local Similarity 94.6%; Pred. No. 2.5e-111;
Matches 264; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY 1 MONHVISLASAABERRAHADTFGRGIPPOFPDAMPSEERLEQMAELVPGISAHPIYLSG 60
DB 1 MONHVISLASAABERRAHADTFGRGIPPOFPDAMPSEERLEQMAELVPGISAHPIYLSG 60
QY 61 VEKACFMSHAVLMKQALDEGLPYITVFEDDVLLGGESEKFLAEDAMLOERPDPTAFIVR 120
DB 61 VEKACFMSHAVLMKQALDEGLPYITVFEDDVLLGGESEKFLAEDAMLOERPDPTAFIVR 120
QY 121 LETMFMAHVLTPSPGVADYCGRAFPFLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
DB 121 LETMFMAHVLTPSPGVADYCGRAFPFLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
QY 121 LETMFMAHVLTPSPGVADYCGRAFPFLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
DB 121 LETMFMAHVLTPSPGVADYCGRAFPFLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
QY 181 VDLMMFSDPFDREGMPVCQALNPALCAQELHYAKFHONSAAGSLIEHDLINRKQQRDS 240
DB 181 VDLMMFSDPFDREGMPVCQALNPALCAQELHYAKFHONSAAGSLIEHDLINRKQQRDS 240
QY 241 PANTFKHRLIRALTISRERERKRRORRQFIVPFQ 275
DB 241 PANTFKHRLIRALTISRERERKRRORRQFIVPFQ 275

RESULT 9
Q7WUX2 NEIPO
ID Q7WUX2 NEIPO PRELIMINARY; PRT; 279 AA.
AC Q7WUX2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE LgtB.
GN Name=1gtB;
OS Neisseria polysacchara.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxId=489;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX Zhu P., Tsai C.M.;
RT "Galactosyltransferase gene family in Neisseria: Implication for
RT understanding the evolution of prokaryotic enzymes responsible for
RT glycosylation.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF134878; AAN08518.1; -; Genomic DNA.
DR EMBL; AF134878; AAN08518.1; -; Genomic DNA.
SQ
SEQUENCE 279 AA; 31907 MW; 60BDARBD749D46D5 CRC64;

Query Match 93.6%; Score 1353; DB 2; Length 279;
Best Local Similarity 93.5%; Pred. No. 1e-110;
Matches 261; Conservative 3; Mismatches 11; Indels 4; Gaps 1;

QY 1 MONHVISLASAABERRAHADTFGRGIPPOFPDAMPSEERLEQMAELVPGISAHPIYLSG 60
DB 1 MONHVISLASAABERRAHADTFGRGIPPOFPDAMPSEERLEQMAELVPGISAHPIYLSG 60
QY 61 VEKACFMSHAVLMKQALDEGLPYITVFEDDVLLGGESEKFLAEDAMLOERPDPTAFIVR 120
DB 61 VEKACFMSHAVLMKQALDEGLPYITVFEDDVLLGGESEKFLAEDAMLOERPDPTAFIVR 120
QY 121 LETMFMAHVLTPSPGVADYCGRAFPFLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
DB 121 LETMFMAHVLTPSPGVADYCGRAFPFLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
QY 121 LETMFMAHVLTPSPGVADYCGRAFPFLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
DB 121 LETMFMAHVLTPSPGVADYCGRAFPFLSESHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
QY 181 VDLMMFSDPFDREGMPVCQALNPALCAQELHYAKFHONSAAGSLIEHDLINRKQQRDS 240

```



```

DB      181 VDMWNGNDPDRGMPVQOLNPALCAQELHYAKFHONSALGSLIHDRLNKKQRRDS 240
QY      241 PANTFGRRLIRALTISRERKRRORREO----FIVEPQ 275
DB      241 PANTFGRRLIRALTISRERKRRORREOIGKTIIVPQ 279

RESULT 10
OBL2V9 NEIME
ID      OBL2V9_NEIME PRELIMINARY; PRT; 279 AA.
AC      OBL2V9;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      LgtB.
GN      Name=lgfB;
OS      Neisseria meningitidis.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OX      Neisseriaceae; Neisseria.
NCBI_TaxID=487;
RN      NUCLEOTIDE SEQUENCE.
RC      STRAIN=35E; PubMed=12055303;
RA      Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT      "Genetic diversity of three lgt loci for biosynthesis of
RL      lipooligosaccharide (LOS) in Neisseria species.";
DR      EMBL; AF470655; AM33850.1; -; Genomic DNA.
DR      GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR      InterPro; IPR002654; Glyco trans 25; I.
DR      Pfam; PF01755; Glyco trans 25; I.
SQ      SEQUENCE 279 AA; 31962 MW; 4321C9A18C44D075 CRC64;

Query Match      92.2%; Score 1333; DB 2; Length 279;
Best Local Similarity 92.1%; Pred. No. 5.9e-109;
Matches 257; Conservative 6; Mismatches 12; Indels 4; Gaps 1;

QY      1 MONHVISLASAARRAHIAIDTFRGRIIPQFPDAMPSERLQAMALVPGLSAHPTLSG 60
DB      1 MONHVISLASAARRAHIAIDTFRGRIIPQFPDAMPSERLQAMALVPGLSAHPTLSG 60
QY      61 VERACFMSHAVLWKQALDGLPYITVFPDDVLLGEGEKFLEADAWLQRRPDTAFIVR 120
DB      61 VERACFMSHAVLWKQALDGLPYITVFPDDVLLGEGEKFLEADAWLQRRPDTAFIVR 120
QY      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAGYIISRKAMRFFLDLRFALPPGGLHP 180
DB      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAGYIISRKAMRFFLDLRFALPPGGLHP 180
QY      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAGYIISRKAMRFFLDLRFALPPGGLHP 180
DB      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAGYIISRKAMRFFLDLRFALPPGGLHP 180
QY      181 VDLMPSPDFDEGMPVQOLNPALCAQELHYAKFHONSALGSLIHDRLNKKQRRDS 240
DB      181 VDLMPSPDFDEGMPVQOLNPALCAQELHYAKFHONSALGSLIHDRLNKKQRRDS 240
QY      241 PANTFGRRLIRALTISRERKRRORREO----FIVEPQ 275
DB      241 PANTFGRRLIRALTISRERKRRORREOIGKTIIVPQ 279

RESULT 11
OBL2V9 NEIME
ID      OBL2V9_NEIME PRELIMINARY; PRT; 275 AA.
AC      OBL2V9;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE      LgtB.
GN      Name=lgfB;
OS      Neisseria meningitidis.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OX      Neisseriaceae; Neisseria.
NCBI_TaxID=487;

```

```

RN      [1]
RC      STRAIN=6275, and M986;
RX      MEDLINE=22051050; PubMed=12055303;
RA      Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT      "Genetic diversity of three lgt loci for biosynthesis of
RL      lipooligosaccharide (LOS) in Neisseria species.";
DR      EMBL; AF470655; AM33853.1; -; Genomic DNA.
DR      EMBL; AF470657; AM33856.1; -; Genomic DNA.
DR      GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR      InterPro; IPR002654; Glyco trans 25.
DR      Pfam; PF01755; Glyco trans 25; I.
SQ      SEQUENCE 275 AA; 31558 MW; A73037B079FF5DB3 CRC64;

Query Match      91.8%; Score 1328; DB 2; Length 275;
Best Local Similarity 92.0%; Pred. No. 1.6e-108;
Matches 253; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY      1 MONHVISLASAARRAHIAIDTFRGRIIPQFPDAMPSERLQAMALVPGLSAHPTLSG 60
DB      1 MONHVISLASAARRAHIAIDTFRGRIIPQFPDAMPSERLQAMALVPGLSAHPTLSG 60
QY      61 VERACFMSHAVLWKQALDGLPYITVFPDDVLLGEGEKFLEADAWLQRRPDTAFIVR 120
DB      61 VERACFMSHAVLWKQALDGLPYITVFPDDVLLGEGEKFLEADAWLQRRPDTAFIVR 120
QY      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAGYIISRKAMRFFLDLRFALPPGGLHP 180
DB      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAGYIISRKAMRFFLDLRFALPPGGLHP 180
QY      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAGYIISRKAMRFFLDLRFALPPGGLHP 180
DB      121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAGYIISRKAMRFFLDLRFALPPGGLHP 180
QY      181 VDLMPSPDFDEGMPVQOLNPALCAQELHYAKFHONSALGSLIHDRLNKKQRRDS 240
DB      181 VDLMPSPDFDEGMPVQOLNPALCAQELHYAKFHONSALGSLIHDRLNKKQRRDS 240
QY      241 PANTFGRRLIRALTISRERKRRORREOIFIVPQ 275
DB      241 PANTFGRRLIRALTISRERKRRORREOIFIVPQ 275

RESULT 12
OBL2V2 NEIME
ID      OBL2V2_NEIME PRELIMINARY; PRT; 275 AA.
AC      OBL2V2;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      LgtB.
GN      Name=lgfB;
OS      Neisseria meningitidis.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OX      Neisseriaceae; Neisseria.
NCBI_TaxID=487;
RN      NUCLEOTIDE SEQUENCE.
RC      STRAIN=6155; PubMed=12055303;
RA      Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT      "Genetic diversity of three lgt loci for biosynthesis of
RL      lipooligosaccharide (LOS) in Neisseria species.";
DR      EMBL; AF470661; AM33867.1; -; Genomic DNA.
DR      GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR      InterPro; IPR002654; Glyco trans 25.
DR      Pfam; PF01755; Glyco trans 25; I.
SQ      SEQUENCE 275 AA; 31623 MW; B435413181C35674 CRC64;

Query Match      91.8%; Score 1327; DB 2; Length 275;
Best Local Similarity 92.4%; Pred. No. 2e-108;
Matches 254; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY      1 MONHVISLASAARRAHIAIDTFRGRIIPQFPDAMPSERLQAMALVPGLSAHPTLSG 60

```

```

Db      1  M0NHV1SLASAEARRAHIAATGAGCIPFOFPDALMPSERLQAMAEVPGLSAHPYLSG 60
Qy      61  VAKACFMSHAVALMKQALDDEGLPYITVFEEDVLLGEGEKEFLAEDAWLQERPPDTAFIVR 120
Db      61  VAKACFMSHAVALMKQALDDEGLPYIAVFEEDVLLGEGEKEFLAEDAWLKERPPDSAFIVR 120
Qy      121  LETMFHVALTSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFFLDLRFALPPEGGLHP 180
Db      121  LETMFHVALTSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFFLDLRFALPPEGGLHP 180
Qy      181  VDLMFSDPFDEBEGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
Db      181  VDLMFSDPFDEBEGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
Qy      241  PANTFKRLIRALTITKISRERKRQRORREOFIVPFQ 275
Db      241  PANTFKRLIRALTITKISRERKRQRORREOFIVPFQ 275

```

RESULT 13

```

QBL2U6_NEIME PRELIMINARY; PRT; 279 AA.
ID      QBL2U6;
AC      QBL2U6;
DT      01-OCT-2002 (T#BMBLrel. 22, Created)
DT      01-OCT-2002 (T#BMBLrel. 22, Last sequence update)
DE      01-JUN-2003 (T#BMBLrel. 24, Last annotation update)
DE      LGTB.
GN      Name=lgTB;
OS      Neisseria meningitidis.
OC      Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC      Neisseriaceae; Neisseria.
OX      NCBI_TaxID=487;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN-7897.
RX      MEDLINE=22051050; PubMed=12055303;
RA      Zhu P., Klutch M.J., Bash M.C., Tsang R.S.W., Ng L.K., Tsai C.M.;
RT      "Genetic diversity of three lgt loci for biosynthesis of
RL      lipooligosaccharide (LOS) in Neisseria species.";
RL      Microbiology 148:1833-1844(2002).
DR      EMBL; AF470665; AM33876.1; -; Genomic DNA.
DR      GO; GO:0009103; P.lipopolysaccharide biosynthesis; IEA.
DR      InterPro; IPR002654; Glyco_trans_25.
DR      Pfam; PF01755; Glyco_transf_25; 1.
SQ      SEQUENCE 279 AA; 31927 MW; C7D3A3BE74ABA3D9 CRC64;

```

```

Query Match      91.6%; Score 1325; DB 2; Length 279;
Best Local Similarity 92.5%; Pred. No. 3e-108;
Matches 258; Conservative 3; Mismatches 14; Indels 4; Gaps 1;

Qy      1  M0NHV1SLASAEARRAHIAATGAGCIPFOFPDALMPSERLQAMAEVPGLSAHPYLSG 60
Db      1  M0NHV1SLASAEARRAHIAATGAGCIPFOFPDALMPSERLQAMAEVPGLSAHPYLSG 60
Qy      61  VAKACFMSHAVALMKQALDDEGLPYITVFEEDVLLGEGEKEFLAEDAWLQERPPDTAFIVR 120
Db      61  VAKACFMSHAVALMKQALDDEGLPYIAVFEEDVLLGEGEKEFLAEDAWLQERPPDSAFIVR 120
Qy      121  LETMFHVALTSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFFLDLRFALPPEGGLHP 180
Db      121  LETMFHVALTSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFFLDLRFALPPEGGLHP 180
Qy      181  VDLMFSDPFDEBEGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
Db      181  VDLMFSDPFDEBEGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
Qy      241  PANTFKRLIRALTITKISRERKRQRORREOFIVPFQ 275
Db      241  PANTFKRLIRALTITKISRERKRQRORREOFIVPFQ 275

```

RESULT 14

LGTB_NEIME

```

ID      LGTB_NEIME STANDARD; PRT; 279 AA.
AC      P57033;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      13-SEP-2005 (Rel. 48, Last annotation update)
DE      Lacto-N-neotetraose biosynthesis glycosyl transferase lgtB
DE      (EC 2.4.1.21).
GN      Name=lgTB; Ordered locus names=MMA0525;
OS      Neisseria meningitidis (serogroup A).
OC      Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC      Neisseriaceae; Neisseria.
OX      NCBI_TaxID=65699;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=22491 / Serotype 4a;
RX      MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA      Parikh I.J., Achman M., James K.D., Bentley S.D., Churcher C.M.,
RA      Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA      Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA      Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA      Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA      Whitehead S., Spratt B.G., Barrall B.G.;
RT      "Complete DNA sequence of a serogroup A strain of Neisseria
RL      meningitidis 22491."
RL      Nature 404:502-506(2000).
CC      -1- FUNCTION: Adds the second galactose to the lacto-N-tetraose chain
CC      in lipooligosaccharide (LOS).
CC      -1- PATHWAY: Biosynthesis of the terminal lacto-N-neotetraose LPS
CC      structure.
CC      -1- SIMILARITY: Belongs to the glycosyltransferase 25 family.
CC      CC
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      CC
CC      EMBL; AL162753; CAB83817.1; -; Genomic DNA.
CC      PIR; A81971; A81971.
DR      InterPro; IPR002654; Glyco_trans_25.
DR      PANTHER; PTHR15862; Glyco_trans_25; 1.
DR      Pfam; PF01755; Glyco_transf_25; 1.
KW      Complete proteome; Glycosyltransferase;
KW      Lipopolysaccharide biosynthesis; Transferase.
SQ      SEQUENCE 279 AA; 31904 MW; 8703B56513A0D347 CRC64;

```

```

Query Match      91.6%; Score 1324; DB 1; Length 279;
Best Local Similarity 92.5%; Pred. No. 3.7e-108;
Matches 258; Conservative 2; Mismatches 15; Indels 4; Gaps 1;

Qy      1  M0NHV1SLASAEARRAHIAATGAGCIPFOFPDALMPSERLQAMAEVPGLSAHPYLSG 60
Db      1  M0NHV1SLASAEARRAHIAATGAGCIPFOFPDALMPSERLQAMAEVPGLSAHPYLSG 60
Qy      61  VAKACFMSHAVALMKQALDDEGLPYITVFEEDVLLGEGEKEFLAEDAWLQERPPDTAFIVR 120
Db      61  VAKACFMSHAVALMKQALDDEGLPYIAVFEEDVLLGEGEKEFLAEDAWLQERPPDSAFIVR 120
Qy      121  LETMFHVALTSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFFLDLRFALPPEGGLHP 180
Db      121  LETMFHVALTSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFFLDLRFALPPEGGLHP 180
Qy      181  VDLMFSDPFDEBEGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
Db      181  VDLMFSDPFDEBEGMPCVQALNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
Qy      241  PANTFKRLIRALTITKISRERKRQRORREOFIVPFQ 275
Db      241  PANTFKRLIRALTITKISRERKRQRORREOFIVPFQ 275

```

RESULT 15

Q93EK5_NEIME

ID Q93EK5_NEIME_PRELIMINARY; PRT; 275 AA.
AC Q93EK5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE LgtB.
GN Name=LgtB;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M978;
RX MEDLINE=21467954; Pubmed=11583844;
RA Zhu P., Klutch M.J., Tsai C.-M.;
RT "Genetic analysis of conservation and variation of lipooligosaccharide
expression in two L8-immunotype strains of Neisseria meningitidis.";
RL FEMS Microbiol. Lett. 203:173-177(2001).
DR EMBL; AF355193; AAL12841.1; -; Genomic DNA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro; IPR002654; Glyco trans 25.
DR Pfam; PF01755; Glyco_transf_25; I.
SQ SEQUENCE 275AA; 31577 MW; A73037E0625EE7B3 CRC64;

Query Match 91.3%; Score 1320; DB 2; Length 275;
Best local similarity 91.6%; Pred. No. 8,1e-108;
Matches 252; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 MGNHVISLSAARRAHIAIDTFSRHGIPQFPDAMPSERLBOAMALVPGLSAHPYLSG 60
DB 1 MGNHVISLSAARRAHIAIDTFSRHGIPQFPDAMPSERLBOAMALVPGLSAHPYLSG 60
QY 61 VERACMSNAVLWKQALDEGLPYITVFEDVLLGEGEKFLEADAWLQERFPDPTAFIVR 120
DB 61 VERACMSNAVLWKQALDEGLPYITVFEDVLLGEGEKFLEADAWLQERFPDPTAFIVR 120
QY 121 LETMFHVLTPSGVADYCGRAFPLLESEHMGTAGYTIISRKAMPFLDRPALPPGLHP 180
DB 121 LETMFHVLTPSGVADYCGRAFPLLESEHMGTAGYTIISRKAMPFLDRPALPPGLHP 180
QY 181 VDLMMFSDFPDREGMPCQALPALCAQELHYAKFDONSALGSLIEHRLNKKQQRDS 240
DB 181 VDLMMFSDFPDREGMPCQALPALCAQELHYAKFDONSALGSLIEHRLNKKQQRDS 240
QY 241 PANTFKRLRLATLTKISREKRRQRREQPTVPFQ 275
DB 241 PANTFKRLRLATLTKISREKRRQRREQPTVPFQ 275

Search completed: April 7, 2006, 15:29:48
Job time : 252 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: April 7, 2006, 15:30:06 ; Search time 47 Seconds
(without alignments)
483.741 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446

Sequence: 1 MGNHYISLASAERRAHIAAD.....ISREKRRQRREQIVPQ 275

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/1aa/5 COMB.pep:*\n2: /cgn2_6/prodata/1/1aa/6 COMB.pep:*\n3: /cgn2_6/prodata/1/1aa/H COMB.pep:*\n4: /cgn2_6/prodata/1/1aa/PCUTUS COMB.pep:*\n5: /cgn2_6/prodata/1/1aa/RB COMB.pep:*\n6: /cgn2_6/prodata/1/1aa/Backfill1seq.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1288	89.1	279	1	US-08-312-387B-8 Sequence 8, Appl
2	1288	89.1	279	1	US-08-683-426-8 Sequence 8, Appl
3	1288	89.1	279	1	US-08-683-458-8 Sequence 8, Appl
4	1288	89.1	279	1	US-08-878-360-8 Sequence 8, Appl
5	1288	89.1	279	2	US-09-333-412-8 Sequence 8, Appl
6	1288	89.1	279	2	US-10-007-267A-8 Sequence 8, Appl
7	1007	69.6	280	1	US-08-312-387B-6 Sequence 6, Appl
8	1007	69.6	280	1	US-08-683-426-6 Sequence 6, Appl
9	1007	69.6	280	1	US-08-683-458-6 Sequence 6, Appl
10	1007	69.6	280	1	US-08-878-360-6 Sequence 6, Appl
11	1007	69.6	280	2	US-08-478-140B-6 Sequence 6, Appl
12	1007	69.6	280	2	US-09-333-412-6 Sequence 6, Appl
13	1007	69.6	280	2	US-09-338-943-6 Sequence 6, Appl
14	1007	69.6	280	2	US-10-007-267A-6 Sequence 6, Appl
15	364	25.2	255	2	US-09-540-236-2407 Sequence 2407, Ap
16	118.5	8.2	579	2	US-09-369-247-61 Sequence 61, Appl
17	118.5	8.2	579	2	US-10-062-548-61 Sequence 61, Appl
18	98.5	6.8	273	2	US-09-601-519A-1 Sequence 1, Appl
19	98.5	6.8	273	2	US-10-218-381-1 Sequence 1, Appl
20	93	6.4	114	2	US-09-344-587-10 Sequence 10, Appl
21	93	6.4	150	1	US-08-400-115-2 Sequence 2, Appl
22	93	6.4	271	1	US-08-400-115-4 Sequence 4, Appl
23	92	6.4	244	2	US-09-244-369B-1 Sequence 1, Appl
24	92	6.4	244	2	US-09-940-391-1 Sequence 1, Appl
25	92	6.4	255	2	US-09-328-352-5253 Sequence 5253, Ap
26	91	6.3	116	2	US-09-244-592-1 Sequence 1, Appl
27	91	6.3	133	2	US-08-961-309-48 Sequence 48, Appl

28	90	6.2	5069	2	US-10-042-665A-5 Sequence 5, Appl
29	89	6.2	133	1	US-08-822-028-6 Sequence 6, Appl
30	89	6.2	133	1	US-08-822-028-30 Sequence 30, Appl
31	89	6.2	133	2	US-08-479-285-6 Sequence 6, Appl
32	89	6.2	133	2	US-08-479-285-30 Sequence 30, Appl
33	89	6.2	133	2	US-09-503-653A-6 Sequence 6, Appl
34	89	6.2	133	2	US-09-503-653A-30 Sequence 30, Appl
35	88.5	6.1	134	2	US-08-961-309-49 Sequence 49, Appl
36	87.5	6.1	238	1	US-08-224-591-12 Sequence 12, Appl
37	87.5	6.1	238	1	US-08-392-338A-21 Sequence 21, Appl
38	87.5	6.1	238	1	US-08-926-789-12 Sequence 12, Appl
39	87.5	6.1	238	2	US-09-166-750-21 Sequence 21, Appl
40	87.5	6.1	238	2	US-09-166-093-21 Sequence 21, Appl
41	87.5	6.1	238	2	US-09-172-019-21 Sequence 21, Appl
42	87.5	6.1	238	2	US-09-166-094-21 Sequence 21, Appl
43	87.5	6.1	238	2	US-09-443-213-21 Sequence 21, Appl
44	87.5	6.1	239	4	PCT-US93-11138-12 Sequence 12, Appl
45	86.5	6.0	134	1	US-08-822-028-10 Sequence 10, Appl

ALIGNMENTS

```
RESULT 1
US-08-312-387B-8
; Sequence 8, Application US/08312387B
; Patent No. 5545553
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,387B
; FILING DATE: July 7, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-312-387B-8
Query Match 89.1%; Score 1288; DB 1; Length 279;
Best Local Similarity 90.0%; Pred. No. 9.9e+143;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;
```

1 MGNHYISLASAERRAHIAADTFGRHGIPPOFDALMPSEBRLQMAAEIVPGISAPHTYISG 60
1 MGNHYISLASAERRAHIAATFGSGRGIPOFDALMPSEBRLERRAAAEIVPGISAPHTYISG 60
61 VEKACFMHSAVLMKQALDEGLPYITVFEDDVILGSGEKKFLAEDAMLDGERPDPAFTYR 120

DB 61 VAKACFMSHVAVLMQALDEGVPIYIAVFEDVLLGSGAQFLAEDTWLQERFPDPSAFVVR 120
QY 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDLRFALPPEGILHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDLRFALPPEGILHP 180
QY 181 VDLMMFSDPFDREGMPVQOLNPALCAQELHYAKFHDQNSALGSLIHDRLNLRKQWRDS 240
DB 181 VDLMMFSDPFDREGMPVQOLNPALCAQELHYAKFHDQNSALGSLIHDRLNLRKQWRDS 240
QY 241 PANTFKHRLIRALTTKISREKRQRROQ---FIVPFO 275
DB 241 PANTFKHRLIRALTTKISREKRQRROQ---FIVPFO 275

RESULT 2
US-08-683-426-8
; Sequence 8, Application US/08683426
; Patent No. 5705367
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,426
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: September 26, 1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULAR TYPE: protein
US-08-683-426-8

Query Match 89.1%; Score 1288; DB 1; Length 279;
Best Local Similarity 90.0%; Pred. No. 9,9e-143;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDLRFALPPEGILHP 180
DB 121 LETMFMHVLTPSPGVADYCGRAFPLLSESHMGTAGYIISRKAMRFFLDLRFALPPEGILHP 180
QY 181 VDLMMFSDPFDREGMPVQOLNPALCAQELHYAKFHDQNSALGSLIHDRLNLRKQWRDS 240
DB 181 VDLMMFSDPFDREGMPVQOLNPALCAQELHYAKFHDQNSALGSLIHDRLNLRKQWRDS 240
QY 241 PANTFKHRLIRALTTKISREKRQRROQ---FIVPFO 275
DB 241 PANTFKHRLIRALTTKISREKRQRROQ---FIVPFO 275

RESULT 3
US-08-683-458-8
; Sequence 8, Application US/08683458
; Patent No. 5798233
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,458
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/312,387
; FILING DATE: September 26, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULAR TYPE: protein
US-08-683-458-8

Query Match 89.1%; Score 1288; DB 1; Length 279;
Best Local Similarity 90.0%; Pred. No. 9,9e-143;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY	181	VDIMAFSEPFDEGMPVCCOLNPAICQOELHYAFHONSALGSLIHDRLNKKOORDS	240
Db	181	VDIMAFSEGNDDEGMPVCCOLNPAICQOELHYAFHONSALGSLIHDRLNKKOORDS	240
QY	241	PANTPKHRLIRALTYKISREERKRRORREO----	275
Db	241	PANTPKHRLIRALTYKISREERKRRORREOLIGKIIVPFO	279

RESULT 4

US-08-878-360-8
: Sequence 8, Application us/08878360
: Patent No. 5945322
: GENERAL INFORMATION:
: APPLICANT: Gotschlich, Emil C.
: TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
: TITLE OF INVENTION: OLIGOSACCHARIDES, AND GENES ENCODING THEM
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauber & Jackson
: STREET: 411 Hackensack Avenue
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/878,360
: FILING DATE: 18-JUN-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/683,426
: FILING DATE:
: APPLICATION NUMBER: 08/312,387
: FILING DATE: September 26, 1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 600-1-095B
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201 487-5800
: TELEFAX: 201 343-1684
: TELEFAX: 133521
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 279 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-878-360-8

Query Match	89.1%;	Score 1288;	DB 1;	Length 279;
Best Local Similarity	90.0%;	Pred. No. 9.9e-143;		
Matches 251; Conservative	7;	Mismatches 17;	Indels 4;	Gaps 1;

Qy	1	MÖNHVYSLASAEREAHIDTFGRGDI	PROFPAALMPSRRLEQAMVELPGSAHPYLS	60
Db	1	MÖNHVYSLASAEREAHIDTFGRGDI	PROFPAALMPSRRLEQAMVELPGSAHPYLS	60
Qy	61	VEKAFEMHAYLTKMALJDEGLPYITWPEDDYLL	IGEEEEKFLPADALQSRPDDTAYIR	120
Db	61	VEKAFEMHAYLTKMALJDEGLPYITWPEDDYLL	IGEEEEKFLPADALQSRPDDTAYIR	120
Qy	61	VEKAFEMHAYLTKMALJDEGLPYITWPEDDYLL	IGEEEEKFLPADALQSRPDDTAYIR	120
Db	61	VEKAFEMHAYLTKMALJDEGLPYITWPEDDYLL	IGEEEEKFLPADALQSRPDDTAYIR	120
Qy	121	LETMMHAYLTSSPGVADYCGRAAPLIESSEHWGTAGT	ISRKAMRFLIDRPAALPPEGLHP	180
Db	121	LETMMHAYLTSSPGVADYCGRAAPLIESSEHWGTAGT	ISRKAMRFLIDRPAALPPEGLHP	180
Qy	121	LETMMHAYLTSSPGVADYCGRAAPLIESSEHWGTAGT	ISRKAMRFLIDRPAALPPEGLHP	180
Db	121	LETMMHAYLTSSPGVADYCGRAAPLIESSEHWGTAGT	ISRKAMRFLIDRPAALPPEGLHP	180

Qy	181	VDLMMESPPFREGPVCQINLPALCAQELHTAKFDONSALGSLIENHDLNLRKQORRS	240
Db	181	VDLMMFGNPDREGPVCQINLPALCAQELHTAKFDONSALGSLIENHDLNLRKQORRS	240
Qy	241	PANTPKRLILALTYKISREKRRKRRRQ----	275
Db	241	PANTPKRLILALTYKISREKRRRQRLIKTIIIVPFO	279

RESULT 5

```

US-09-333-412-8
/ Sequence 8, Application US/09333412
/ Patent No. 6342382
/ GENERAL INFORMATION:
/ APPLICANT: Gotschlich, Emil C.
/ TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
/ OLIGOSACCHARIDES, AND GENES ENCODING THEM
/
/ NUMBER OF SEQUENCES: 12
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Klauber & Jackson
/ STREET: 411 Hackensack Avenue
/ CITY: Hackensack
/ STATE: New Jersey
/ COUNTRY: USA
/ ZIP: 07601
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/333,412
/ FILING DATE: 15-Jun-1999
/
/ CLASSIFICATION: <Unknown>
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/312,387
/ FILING DATE: July 7, 1994
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Jackson Esq., David A.
/ REGISTRATION NUMBER: 26,742
/ REFERENCE/DOCKET NUMBER: 600-1-095
/
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 201 487-5600
/ TELEFAX: 201 343-1684
/ TELEX: 133521
/
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 279 amino acids
/ TYPE: amino acid
/ TOPOLOGY: 1linear
/
/ MOLECULE TYPE: protein
/
/ SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-333-412-8

```

Query Match	89.1%;	Score 1288;	DB 2;	Length 279;
Best Local Similarity	90.0%;	Pred. NO. 9.9e-143;		
Matches 251; Conservative	7;	Mismatches 17;	Indels 4;	Gaps 1;

Qy	1	MONHVI	SLAASAE	RAH	AD	FE	GH	GI	P	EO	F	D	L	M	P	S	E	R	L	E	O	A	A	I	V	P	G	S	A	P	T	Y	S	G	60																			
Db	1	MONHVI	SLAASAE	RAH	AD	FE	GH	GI	P	EO	F	D	L	M	P	S	E	R	L	E	O	A	A	I	V	P	G	S	A	P	T	Y	S	G	60																			
Qy	61	VEKAC	FMSH	AV	LT	M	R	Q	A	L	D	E	G	L	P	Y	I	T	V	E	P	D	D	V	L	G	E	B	E	K	F	L	A	E	D	A	M	L	O	R	P	D	T	A	V	I	R	120						
Db	61	VEKAC	FMSH	AV	LT	M	R	Q	A	L	D	E	G	L	P	Y	I	T	V	E	P	D	D	V	L	G	E	B	E	K	F	L	A	E	D	A	M	L	O	R	P	D	T	A	V	I	R	120						
Qy	121	LET	M	F	M	H	V	L	T	S	P	S	G	V	A	D	Y	C	G	R	A	P	L	L	B	E	S	H	M	G	T	A	G	Y	I	I	S	R	K	A	M	F	L	D	R	P	A	L	P	E	G	L	H	180
Db	121	LET	M	F	M	H	V	L	T	S	P	S	G	V	A	D	Y	C	G	R	A	P	L	L	B	E	S	H	M	G	T	A	G	Y	I	I	S	R	K	A	M	F	L	D	R	P	A	L	P	E	G	L	H	180
Qy	181	VD	L	M	F	M	E	D	F	P	R	E	G	M	P	V	Q	L	N	P	A	C	O	E	L	H	A	K	F	H	D	O	N	S	A	L	G	S	L	I	E	H	R	L	N	R	K	O	O	R	D	S	240	
Db	181	VD	L	M	F	M	E	D	F	P	R	E	G	M	P	V	Q	L	N	P	A	C	O	E	L	H	A	K	F	H	D	O	N	S	A	L	G	S	L	I	E	H	R	L	N	R	K	O	O	R	D	S	240	

APPLICATION NUMBER: US/08/683,426
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-426-6

Query Match 69.6%; Score 1007; DB 1; Length 280;
Best Local Similarity 74.7%; Pred. No. 1.1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MONHVSLASAERRAHIDTGRHGIPQFPDAMPSERLQAMAEVPGLSAHPYLSG 60
DB 1 MONHVSLASAERRAHIDTGRHGIPQFPDAMPSERLQAMAEVPGLSAHPYLSG 60
QY 61 VERACMSHAIVLWKQALDGLPYITVFEDVLLGEGEKEFLADAVLQERFPDPTAFYR 120
DB 61 VERACMSHAIVLWKQALDGLPYITVFEDVLLGEGEKEFLADAVLQERFPDPTAFYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPPLISESHWGTAGYIISRKARFPLDRFALPPEGGLP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLISESHWGTAGYIISRKARFPLDRFALPPEGGLP 180
QY 121 LETMFHVLTPSPGVADYCGRAFPPLISESHWGTAGYIISRKARFPLDRFALPPEGGLP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLISESHWGTAGYIISRKARFPLDRFALPPEGGLP 180
QY 181 VDLMPSPDFRDEGMPVCOINPALCAQELHYAKFDONSALSLIHHDLINKQQRDS 240
DB 181 VDLMPSPDFRDEGMPVCOINPALCAQELHYAKFDONSALSLIHHDLINKQQRDS 240
QY 181 VDLMPSPDFRDEGMPVCOINPALCAQELHYAKFDONSALSLIHHDLINKQQRDS 240
DB 181 VDLMPSPDFRDEGMPVCOINPALCAQELHYAKFDONSALSLIHHDLINKQQRDS 240
QY 241 PANTFHKRLIRALTKISREKRRORRQ 269
DB 241 PANTFHKRLIRALTKISREKRRORRQ 269
QY 237 RSLKVMFDLKRALGKFKREKKRMRORQ 265
DB 237 RSLKVMFDLKRALGKFKREKKRMRORQ 265

RESULT 9
US-08-683-458-6
Sequence 6, Application US/08683458
Patent No. 5798233
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSES: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/683,458
APPLICATION NUMBER: 435
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-458-6

Query Match 69.6%; Score 1007; DB 1; Length 280;
Best Local Similarity 74.7%; Pred. No. 1.1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MONHVSLASAERRAHIDTGRHGIPQFPDAMPSERLQAMAEVPGLSAHPYLSG 60
DB 1 MONHVSLASAERRAHIDTGRHGIPQFPDAMPSERLQAMAEVPGLSAHPYLSG 60
QY 61 VERACMSHAIVLWKQALDGLPYITVFEDVLLGEGEKEFLADAVLQERFPDPTAFYR 120
DB 61 VERACMSHAIVLWKQALDGLPYITVFEDVLLGEGEKEFLADAVLQERFPDPTAFYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPPLISESHWGTAGYIISRKARFPLDRFALPPEGGLP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLISESHWGTAGYIISRKARFPLDRFALPPEGGLP 180
QY 121 LETMFHVLTPSPGVADYCGRAFPPLISESHWGTAGYIISRKARFPLDRFALPPEGGLP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLISESHWGTAGYIISRKARFPLDRFALPPEGGLP 180
QY 181 VDLMPSPDFRDEGMPVCOINPALCAQELHYAKFDONSALSLIHHDLINKQQRDS 240
DB 181 VDLMPSPDFRDEGMPVCOINPALCAQELHYAKFDONSALSLIHHDLINKQQRDS 240
QY 181 VDLMPSPDFRDEGMPVCOINPALCAQELHYAKFDONSALSLIHHDLINKQQRDS 240
DB 181 VDLMPSPDFRDEGMPVCOINPALCAQELHYAKFDONSALSLIHHDLINKQQRDS 240
QY 241 PANTFHKRLIRALTKISREKRRORRQ 269
DB 241 PANTFHKRLIRALTKISREKRRORRQ 269
QY 237 RSLKVMFDLKRALGKFKREKKRMRORQ 265
DB 237 RSLKVMFDLKRALGKFKREKKRMRORQ 265

RESULT 10
US-08-878-360-6
Sequence 6, Application US/08878360
Patent No. 5945322
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSES: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/878,360
APPLICATION NUMBER: 435
PRIOR APPLICATION DATA:
FILING DATE: 08/683,426

APPLICATION NUMBER: 08/312,387
FILING DATE: September 26, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-0958
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-360-6

Query Match 69.6%; Score 1007; DB 1; Length 280;
Best Local Similarity 74.7%; Pred. No. 1.1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MONHVISLASAARRAHADTFGRGIPFOFPDAMPSERLQAMAEIVPGISAHPTLSG 60
DB 1 MONHVISLASAARRAHADTFGRGIPFOFPDAMPSERLQAMAEIVPGISAHPTLSG 60
QY 61 VEKACFMSHAYVMKQALDEGLPYITVFEDDVLLGEGEKFLEADAWLOERPDPTAFIVR 120
DB 61 VEKACFMSHAYVMKQALDEGLPYITVFEDDVLLGEGEKFLEADAWLOERPDPTAFIVR 120
QY 121 LETMFHVLTSPSGADYCGRAFPFLSESHNGTAGYIISRKAMRFFLDRFALPPEGLHP 180
DB 121 LETMFHVLTSPSGADYCGRAFPFLSESHNGTAGYIISRKAMRFFLDRFALPPEGLHP 180
QY 181 VDLMMSPDFDEBEGMPVQALNPALCAQELHYAKFHQNSALGSLIEHRLNRKQORSDS 240
DB 181 VDLMMFTYFDEKEMPVYQVSPALCTQELHYAKFLSQNSMLGSDLEKD----REGQRRRR 236
QY 241 PANTFKHRLIRALTISRERERORRQ 269
DB 237 RSLKVMFDLKRALGKGRKKRMRORQ 265

RESULT 11

US-08-478-140B-6
Sequence 6, Application US/08478140B
Patent No. 6127153
GENERAL INFORMATION:
APPLICANT: JOHNSON, KARL F.
APPLICANT: ROTH, STEPHEN
APPLICANT: BUCZALA, STEPHANIE L.
TITLE OF INVENTION: METHOD OF TRANSPERRING AT LEAST TWO
TITLE OF INVENTION: SACHARIDE UNITS WITH A POLYGLYCOSYLTRANSFERASE, A
TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE AND GENE ENCODING A
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,140B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Laura A. Coruzzi
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7188-017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-140B-6

Query Match 69.6%; Score 1007; DB 2; Length 280;
Best Local Similarity 74.7%; Pred. No. 1.1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MONHVISLASAARRAHADTFGRGIPFOFPDAMPSERLQAMAEIVPGISAHPTLSG 60
DB 1 MONHVISLASAARRAHADTFGRGIPFOFPDAMPSERLQAMAEIVPGISAHPTLSG 60
QY 61 VEKACFMSHAYVMKQALDEGLPYITVFEDDVLLGEGEKFLEADAWLOERPDPTAFIVR 120
DB 61 VEKACFMSHAYVMKQALDEGLPYITVFEDDVLLGEGEKFLEADAWLOERPDPTAFIVR 120
QY 121 LETMFHVLTSPSGADYCGRAFPFLSESHNGTAGYIISRKAMRFFLDRFALPPEGLHP 180
DB 121 LETMFHVLTSPSGADYCGRAFPFLSESHNGTAGYIISRKAMRFFLDRFALPPEGLHP 180
QY 181 VDLMMSPDFDEBEGMPVQALNPALCAQELHYAKFHQNSALGSLIEHRLNRKQORSDS 240
DB 181 VDLMMFTYFDEKEMPVYQVSPALCTQELHYAKFLSQNSMLGSDLEKD----REGQRRRR 236
QY 241 PANTFKHRLIRALTISRERERORRQ 269
DB 237 RSLKVMFDLKRALGKGRKKRMRORQ 265

RESULT 12

US-09-333-412-6
Sequence 6, Application US/09333412
Patent No. 6342382
GENERAL INFORMATION:
APPLICANT: Gotschlich, Em1 C
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
OLIGOSACCHARIDES, AND GENES ENCODING THEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/333,412
FILING DATE: 15-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/312,387
FILING DATE: July 7, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-095

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-333-412-6

Query Match 69.6%; Score 1007; DB 2; Length 280;
Best Local Similarity 74.7%; Pred. No. 1,1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MGNHVISLSAASARRAHIAIDTFSRHGIPQFPDAMPSERLBOAMAEVPGLSAHPYLSG 60
DB 1 MGNHVISLSAASARRAHIAIDTFSRGIPQFPDAMPSERLBOAMAEVPGLSAHPYLSG 60
QY 61 VERACFMSHVAIWKQALDEGLPYITVFEEDVLLGEGEKKFLADAWLQERPDPTAFYR 120
DB 61 VERACFMSHVAIWKQALDEGLPYIAVFEEDVLLGEGAEQFLADPTWLEERFDKSAFIYR 120
QY 121 LETMFHNVLTSPGVDYCGRAFPILSEHMGTAGYIISRKARFPLDRPALPPEGILAP 180
DB 121 LETMFHNVLTSPGVDYCGRAFPILSEHMGTAGYIISRKARFPLDRPALPPEGILAP 180
QY 181 VDLMEFSDPFDEBGMVPCQINPALCAQELHYAKFHONSALGSLIEHDLNLRKQORSD 240
DB 181 VDLMEFSDPFDEBGMVPCQINPALCAQELHYAKFHONSALGSLIEHDLNLRKQORSD 240
QY 241 PANTFKHRLIRALTYISREKRRORREQ 269
DB 241 PANTFKHRLIRALTYISREKRRORREQ 269
QY 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265
DB 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265

RESULT 13

US-09-338-943-6
Sequence 6, Application US/09338943
Patent No. 6379933
GENERAL INFORMATION:
APPLICANT: JOHNSON, KARL F.
APPLICANT: BOTH, STEPHEN
APPLICANT: BUCZALA, STEPHANIE L.
TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO
TITLE OF INVENTION: SACHARIDE UNITS WITH A POLYGLYCOSYLTRANSFERASE, A
TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE AND GENE ENCODING A
TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/338,943
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,140
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Laura A. Cortuzzi
REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7188-017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-338-943-6

Query Match 69.6%; Score 1007; DB 2; Length 280;
Best Local Similarity 74.7%; Pred. No. 1,1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MGNHVISLSAASARRAHIAIDTFSRHGIPQFPDAMPSERLBOAMAEVPGLSAHPYLSG 60
DB 1 MGNHVISLSAASARRAHIAIDTFSRGIPQFPDAMPSERLBOAMAEVPGLSAHPYLSG 60
QY 61 VERACFMSHVAIWKQALDEGLPYITVFEEDVLLGEGEKKFLADAWLQERPDPTAFYR 120
DB 61 VERACFMSHVAIWKQALDEGLPYIAVFEEDVLLGEGAEQFLADPTWLEERFDKSAFIYR 120
QY 121 LETMFHNVLTSPGVDYCGRAFPILSEHMGTAGYIISRKARFPLDRPALPPEGILAP 180
DB 121 LETMFHNVLTSPGVDYCGRAFPILSEHMGTAGYIISRKARFPLDRPALPPEGILAP 180
QY 181 VDLMEFSDPFDEBGMVPCQINPALCAQELHYAKFHONSALGSLIEHDLNLRKQORSD 240
DB 181 VDLMEFSDPFDEBGMVPCQINPALCAQELHYAKFHONSALGSLIEHDLNLRKQORSD 240
QY 241 PANTFKHRLIRALTYISREKRRORREQ 269
DB 241 PANTFKHRLIRALTYISREKRRORREQ 269
QY 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265
DB 237 RSLKVMFDLKRALGKFGREKKRMRORQ 265

RESULT 14

US-10-007-267A-6
Sequence 6, Application US/10007267A
Patent No. 6780624
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
FILE REFERENCE: 040853-01-5029-02
CURRENT APPLICATION NUMBER: US/10/007,267A
CURRENT FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: US 09/333,412
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: US 08/878,360
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: US 08/683,426
PRIOR FILING DATE: 1996-07-18
PRIOR APPLICATION NUMBER: US 08/312,387
PRIOR FILING DATE: 1994-09-24
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.2
SEQ ID NO 6
LENGTH: 280
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-007-267A-6

Query Match 69.6%; Score 1007; DB 2; Length 280;
Best Local Similarity 74.7%; Pred. No. 1,1e-109;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MGNHVISLSAASARRAHIAIDTFSRHGIPQFPDAMPSERLBOAMAEVPGLSAHPYLSG 60
DB 1 MGNHVISLSAASARRAHIAIDTFSRGIPQFPDAMPSERLBOAMAEVPGLSAHPYLSG 60

```

QY 61 VEKACFMSHVAVLWKQALDEGLPYITVFPEDDVLGSGEEKFLAEDAMLOERFPDPTAFIV 120
DB 61 VEKACFMSHVAVLWKQALDEGLPYITVFPEDDVLGSGAEQFLAEDTWLERFPDQSAFIVR 120
QY 121 LETMFMHVLITSPSGVADYCGRAFPILSESHMGCTAGYIISRKAMRPFILDRFPAALPREGIHP 180
DB 121 LETMFMHVLITSPSGVADYCGRAFPILSESHMGCTAGYIISRKAMRPFILDRFPAALPREGIHP 180
QY 181 VDLMMFSPDFDEGMPVQCOLNPALCAOEIHYAKFPHDONSALGSLIEHDRLINRKOQRDS 240
DB 181 VDLMMFSPDFDEGMPVQCOLNPALCAOEIHYAKFPHDONSALGSLIEHDRLINRKOQRDS 240
QY 241 PANTFKHRLIRALTKISRRERKRQRBRQ 269
DB 237 RSLKVMFDLKRALGKFGREKKRMRERQ 265

```

```

RESULT 15
US-09-540-236-2407
; Sequence 2407, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2407
; LENGTH: 255
; TYPE: PRT
; ORGANISM: M. catarrhalis
US-09-540-236-2407

```

```

Query Match 25.2%; Score 364; DB 2; Length 255;
Best Local Similarity 35.2%; Pred. No. 3.8e-34;
Matches 90; Conservative 49; Mismatches 103; Indels 14; Gaps 7;

QY 1 MGNHYISLASAARRAHITDTRGRGIPQFPDMLMPSERLQMAELVPGLSAHPIYLSG 60
DB 3 IONFVISTATKRRRHHIMCEFGKQGIAPFPDVAITPTDISRYAQKLSITPINORLTDG 62
QY 61 VEKACFMSHVAVLWKQALDEGLPYITVFPEDDVLGSGEEKFLAEDAMLOERFPDPTAFIV 119
DB 63 -EKACFISHVALWQWIDENLDYMAIFEDDVLIGDNDQQFLQELRTWLQ---NAVDT 117
QY 120 RLETFMFMHVLITSPSGVADYCGRAFPILSESHMGCTAGYIISRKAMRPFILDRFPAALPREGIHP 179
DB 118 KLET-WVERKIHKKAVTVLNHRQLCPKTFHTGTAGYVISOQAKIILDYISTLDAFEFP 176
QY 180 PVDLMMFSPDFDEGMPVQCOLNPALCAOEIHYAKFPHDONSALGSLIEHDRLINRKOQRDS 239
DB 177 PIDHVIIDALISK--MSVQVNPAMVIQ---AHLVSEDDTFKSLIETORKQNVNQHRRR 230
QY 240 SPANTFKHRLIRALTK 255
DB 231 TLAD-YGKKYRSIGK 245

```

Search completed: April 7, 2006, 15:31:27
Job time : 49 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:41:34 ; Search time 168 Seconds
(Without alignments)
683.947 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446
Sequence: 1 MGNHVISLSAARRAHAD.....ISRERKRQRQRFIVPQ 275

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications AA Main:
1: /cgn2_6/prodata/1/pubppaa/us07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppaa/us08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubppaa/us09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppaa/us10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppaa/us10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubppaa/us11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1446	100.0	275	3	US-09-211-691-2
2	1446	100.0	275	4	US-10-317-773-2
3	1446	100.0	275	4	US-10-317-428-2
4	1288	89.1	279	4	US-10-007-267-8
5	1288	89.1	279	4	US-10-654-528-8
6	1007	69.6	280	4	US-10-007-267-6
7	1007	69.6	280	4	US-10-096-129-6
8	1007	69.6	280	4	US-10-654-528-6
9	933	64.5	280	5	US-10-472-260-176
10	130.5	9.0	622	4	US-10-398-037-1
11	128.5	8.9	332	4	US-10-335-977-8513
12	122.5	8.5	1128	5	US-11-097-143-15177
13	121	8.4	738	5	US-10-723-860-1579
14	118.5	8.2	423	4	US-10-264-237-1829
15	118.5	8.2	560	5	US-10-723-860-4111
16	118.5	8.2	579	4	US-10-062-548-61
17	118.5	8.2	579	5	US-10-918-745-61
18	118.5	8.2	579	6	US-11-002-755-61
19	118.5	8.2	579	6	US-11-002-755-61
20	118.5	8.2	636	4	US-10-264-049-3148
21	110	7.6	201	4	US-10-335-977-8512
22	110	7.6	739	4	US-10-243-552-865
23	110	7.6	739	4	US-10-276-774-1787
24	104	7.2	762	6	US-11-035-599-28
25	100.5	7.0	168	5	US-10-450-763-52449
26	98.5	6.8	273	4	US-10-218-381-1
27	98	6.8	481	4	US-10-475-587-1

28	97.5	6.7	133	4	US-10-335-977-8517	Sequence 8517, App
29	97.5	6.7	144	4	US-10-335-977-8518	Sequence 8518, App
30	93	6.4	109	4	US-10-251-085B-241	Sequence 241, App
31	93	6.4	109	4	US-10-737-252-241	Sequence 241, App
32	93	6.4	113	5	US-10-700-632-78	Sequence 78, App1
33	93	6.4	150	3	US-09-226-157-2	Sequence 2, App1
34	93	6.4	237	4	US-10-336-210-2	Sequence 2, App1
35	93	6.4	237	4	US-10-336-210-3	Sequence 3, App1
36	93	6.4	211	3	US-09-226-157-4	Sequence 4, App1
37	93	6.4	271	4	US-10-336-210-6	Sequence 6, App1
38	93	6.4	272	4	US-10-336-210-7	Sequence 7, App1
39	92	6.4	244	3	US-09-940-391-1	Sequence 1, App1
40	92	6.4	244	4	US-10-336-210-8	Sequence 8, App1
41	92	6.4	483	5	US-10-858-367-17	Sequence 17, App1
42	91	6.3	133	4	US-10-255-478-48	Sequence 48, App1
43	91	6.3	242	4	US-10-336-210-9	Sequence 9, App1
44	90	6.2	113	4	US-10-251-085B-235	Sequence 235, App
45	90	6.2	113	4	US-10-737-252-235	Sequence 235, App

ALIGNMENTS

```
RESULT 1
US-09-211-691-2
; Sequence 2, Application US/09211691
; Patent No. US20020034805A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Makarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Fusion Proteins for Use in Enzymatic Synthesis of
; FILE REFERENCE: 019957-012910US
; CURRENT APPLICATION NUMBER: US/09/211,691
; CURRENT FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 275
; TYPR: PRT
; ORGANISM: Neisseria meningitidis
US-09-211-691-2

Query Match      100.0%; Score 1446; DB 3; Length 275;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGNHVISLSAARRAHADTFGRGIGPQFPDAMPSSRLROAMAEVPGLSAPPYLSG 60
   |||||||
Db 1 MGNHVISLSAARRAHADTFGRGIGPQFPDAMPSSRLROAMAEVPGLSAPPYLSG 60

QY 61 VKACFMGSHAVIMKQALDEGLPIYTFVBDVLLSGEKKFLAEDAWLQRPDPDTAIVR 120
   |||||||
Db 61 VKACFMGSHAVIMKQALDEGLPIYTFVBDVLLSGEKKFLAEDAWLQRPDPDTAIVR 120

QY 121 LETMFMHVLTSQSGVADYCGRAFPILLESBHWGTAGYIISRKAMRFFLDRFALPPEGGLP 180
   |||||||
Db 121 LETMFMHVLTSQSGVADYCGRAFPILLESBHWGTAGYIISRKAMRFFLDRFALPPEGGLP 180

QY 181 VDLAMFSDPFDREGMPVQOLNPAACAELHYAKFDONSALGSLTEHRLINRKOQRDS 240
   |||||||
Db 181 VDLAMFSDPFDREGMPVQOLNPAACAELHYAKFDONSALGSLTEHRLINRKOQRDS 240

QY 241 PANTFKHLIRALTITISREKRQRORRQOFIVPQ 275
   |||||||
Db 241 PANTFKHLIRALTITISREKRQRORRQOFIVPQ 275

RESULT 2
```

US-10-317-773-2
; Sequence 2, Application US/10317773
; Publication No. US20030180928A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Fusion Protein Comprising a UDP-GalNac 4' Epimerase and a
; FILE REFERENCE: 019633-00081205
; CURRENT APPLICATION NUMBER: US/10/317,773
; PRIOR FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 09/211,691
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-317-773-2

Query Match 100.0%; Score 1446; DB 4; Length 275;
Best Local Similarity 100.0%; Pred. No. 3.1e-141;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNHVISLASAARRAHADTFGRGIPFOFDALMPSERLQMAELVPGISAHPIYLSG 60
DB 1 MNNHVISLASAARRAHADTFGRGIPFOFDALMPSERLQMAELVPGISAHPIYLSG 60
QY 61 VEKACFMSHAYLWKQALDEGLPYITVFEDVDVLLGSESEKFLAEDAWLQERPDPTAFIVR 120
DB 61 VEKACFMSHAYLWKQALDEGLPYITVFEDVDVLLGSESEKFLAEDAWLQERPDPTAFIVR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPILSESHWGTAIISRKAMRFFLDRLNPKQRRDS 180
DB 121 LETMFHVLTPSPGVADYCGRAFPILSESHWGTAIISRKAMRFFLDRLNPKQRRDS 180
QY 181 VDLMPSPDFPREGMPVQNLNPAALCAQELHYAKFHDQNSALGSLIEHDLNLRKQRRDS 240
DB 181 VDLMPSPDFPREGMPVQNLNPAALCAQELHYAKFHDQNSALGSLIEHDLNLRKQRRDS 240
QY 241 PANTFKHRLIRALTIKISREKRRQRREQFIVPFQ 275
DB 241 PANTFKHRLIRALTIKISREKRRQRREQFIVPFQ 275

RESULT 3
US-10-317-428-2 *
; Sequence 2, Application US/10317428
; Publication No. US2003018641A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Young, N. Martin
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Nucleic Acid That Encodes a Fusion Protein
; FILE REFERENCE: 019633-00081105
; CURRENT APPLICATION NUMBER: US/10/317,428
; PRIOR FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/069,443
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 09/211,691
; PRIOR FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Neisseria meningitidis

US-10-317-428-2

Query Match 100.0%; Score 1446; DB 4; Length 275;
Best Local Similarity 100.0%; Pred. No. 3.1e-141;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNHVISLASAARRAHADTFGRGIPFOFDALMPSERLQMAELVPGISAHPIYLSG 60
DB 1 MNNHVISLASAARRAHADTFGRGIPFOFDALMPSERLQMAELVPGISAHPIYLSG 60
QY 61 VEKACFMSHAYLWKQALDEGLPYITVFEDVDVLLGSESEKFLAEDAWLQERPDPTAFIVR 120
DB 61 VEKACFMSHAYLWKQALDEGLPYITVFEDVDVLLGSESEKFLAEDAWLQERPDPTAFIVR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPILSESHWGTAIISRKAMRFFLDRLNPKQRRDS 180
DB 121 LETMFHVLTPSPGVADYCGRAFPILSESHWGTAIISRKAMRFFLDRLNPKQRRDS 180
QY 181 VDLMPSPDFPREGMPVQNLNPAALCAQELHYAKFHDQNSALGSLIEHDLNLRKQRRDS 240
DB 181 VDLMPSPDFPREGMPVQNLNPAALCAQELHYAKFHDQNSALGSLIEHDLNLRKQRRDS 240
QY 241 PANTFKHRLIRALTIKISREKRRQRREQFIVPFQ 275
DB 241 PANTFKHRLIRALTIKISREKRRQRREQFIVPFQ 275

RESULT 4
US-10-007-267-8
; Sequence 8, Application US/10007267
; Publication No. US20020127682A1
; GENERAL INFORMATION:
; APPLICANT: Gettschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: Klauder & Jackson
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/007,267
; FILING DATE: 03-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,412
; FILING DATE: 15-Jun-1999
; APPLICATION NUMBER: 08/312,387
; FILING DATE: July 7, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEFAX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-007-267-8

Query Match 89.1%; Score 1288; DB 4; Length 279;
Best Local Similarity 90.0%; Pred. No. 7,8e-125;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY 1 MONHVISLASAERRAHIAIDTFGRHGI PFOFPDAMPSERLERQMAELVPGLSAHPYLSG 60
DB 1 MONHVISLASAERRAHIAIDTFGRHGI PFOFPDAMPSERLERQMAELVPGLSAHPYLSG 60
QY 61 VERACFMSHAVALMEQALDEGVPIYAVFEDVLLGSEAEQFLADBTWLQERFPDQSAFYVR 120
DB 61 VERACFMSHAVALMEQALDEGVPIYAVFEDVLLGSEAEQFLADBTWLQERFPDQSAFYVR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHMGTAAGIISRKAMRFFLDRFALPPEGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHMGTAAGIISRKAMRFFLDRFALPPEGLHP 180
QY 181 VDLMMFSDPFDEGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDLMMFSDPFDEGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
QY 241 PANTFGRHLIRALTISRERKRRORRQ----FIVPFQ 275
DB 241 PANTFGRHLIRALTISRERKRRORRQ----FIVPFQ 275

RESULT 5
US-10-654-528-8
; Sequence 8, Application US/10654528
; Publication No. US20040043464A1
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: 040853-01-5023-03
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 10/007,267
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 09/333,412
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: US 08/878,360
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: US 08/683,426
; PRIOR FILING DATE: 1996-07-18
; PRIOR APPLICATION NUMBER: US 08/312,387
; PRIOR FILING DATE: 1994-09-24
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 8
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-654-528-8

Query Match 89.1%; Score 1288; DB 4; Length 279;
Best Local Similarity 90.0%; Pred. No. 7,8e-125;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY 1 MONHVISLASAERRAHIAIDTFGRHGI PFOFPDAMPSERLERQMAELVPGLSAHPYLSG 60
DB 1 MONHVISLASAERRAHIAIDTFGRHGI PFOFPDAMPSERLERQMAELVPGLSAHPYLSG 60
QY 61 VERACFMSHAVALMEQALDEGVPIYAVFEDVLLGSEAEQFLADBTWLQERFPDQSAFYVR 120
DB 61 VERACFMSHAVALMEQALDEGVPIYAVFEDVLLGSEAEQFLADBTWLQERFPDQSAFYVR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHMGTAAGIISRKAMRFFLDRFALPPEGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHMGTAAGIISRKAMRFFLDRFALPPEGLHP 180
QY 181 VDLMMFSDPFDEGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDLMMFSDPFDEGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240

DB 181 VDLMMFSDPFDEGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
QY 241 PANTFGRHLIRALTISRERKRRORRQ----FIVPFQ 275
DB 241 PANTFGRHLIRALTISRERKRRORRQ----FIVPFQ 275

RESULT 6
US-10-007-267-6
; Sequence 6, Application US/10007267
; Publication No. US20020127682A1
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF
; OLIGOSACCHARIDES, AND GENES ENCODING THEM
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/007,267
; FILING DATE: 03-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,412
; FILING DATE: 15-Jun-1999
; APPLICATION NUMBER: 08/312,387
; FILING DATE: July 7, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-095
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEK: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-007-267-6

Query Match 69.6%; Score 1007; DB 4; Length 280;
Best Local Similarity 74.7%; Pred. No. 1,1e-95;
Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

QY 1 MONHVISLASAERRAHIAIDTFGRHGI PFOFPDAMPSERLERQMAELVPGLSAHPYLSG 60
DB 1 MONHVISLASAERRAHIAIDTFGRHGI PFOFPDAMPSERLERQMAELVPGLSAHPYLSG 60
QY 61 VERACFMSHAVALMEQALDEGVPIYAVFEDVLLGSEAEQFLADBTWLQERFPDQSAFYVR 120
DB 61 VERACFMSHAVALMEQALDEGVPIYAVFEDVLLGSEAEQFLADBTWLQERFPDQSAFYVR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHMGTAAGIISRKAMRFFLDRFALPPEGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLLESEHMGTAAGIISRKAMRFFLDRFALPPEGLHP 180
QY 181 VDLMMFSDPFDEGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDLMMFSDPFDEGMPVQOLNPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240

Qy 241 PANTFKHRLIRALTKISRERKRRORQ 269
 Db 237 RSLKVMFDLKRALGKFKREKKRMRORQ 265

RESULT 7

US-10-096-129-6
 ; Sequence 6, Application US/10096129
 ; Publication No. US20030207406A1
 ; GENERAL INFORMATION:
 ; APPLICANT: JOHNSON, KARL F.
 ; APPLICANT: ROTH, STEPHEN
 ; APPLICANT: BUCCALA, STEPHANIE L.
 ; TITLE OF INVENTION: METHOD OF TRANSFERRING AT LEAST TWO SACCHARIDE UNITS WITH A
 ; TITLE OF INVENTION: POLYGLYCOSYLTRANSFERASE
 ; FILE REFERENCE: 040853-01-5021-02
 ; CURRENT APPLICATION NUMBER: US/10/096,129
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/338,943
 ; PRIOR FILING DATE: 1999-06-24
 ; PRIOR APPLICATION NUMBER: US 08/478,140
 ; PRIOR FILING DATE: 1995-06-07
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 280
 ; TYPE: PRF
 ; ORGANISM: Neisseria gonorrhoeae
 US-10-096-129-6

Query Match 69.6%; Score 1007; DB 4; Length 280;
 Best Local Similarity 74.7%; Pred. No. 1.1e-95;

Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

Qy 1 MONHVISLASAERRAHIAIDTFGRHGIPOFPDMLPSERLQAMAEIVPGLSAHPYLSG 60
 Db 1 MONHVISLASAERRAHIAIDTFGRSGIPPOFPDMLPSERLQAMAEIVPGLSAHPYLSG 60
 Qy 61 VEKACFMSHVAVLWKQALDGLPYITVFEEDVLLGGESEKFLAEDAMLOERPDPTAFIVR 120
 Db 61 VEKACFMSHVAVLWKQALDGLPYIAVFEEDVLLGGESEKFLAEDTWLERFDKDSAFIVR 120
 Qy 121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAIIISRKAMRFFLDRFALPPEGILHP 180
 Db 121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAIIISRKAMRFFLDRFALPPEGILHP 180
 Qy 181 VDLMPFSDPFDREGMPVQCLNPALCAQELHYAKFHDONSALGSLIENDRLNKKOQRDS 240
 Db 181 VDLMPFSDPFDREGMPVQSPALCTQELHYAKFSLSONSLGSDLEKD---REGRRHR 236
 Qy 241 PANTFKHRLIRALTKISRERKRRORQ 269
 Db 237 RSLKVMFDLKRALGKFKREKKRMRORQ 265

RESULT 8

US-10-654-528-6
 ; Sequence 6, Application US/10654528
 ; Publication No. US20040043464A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gotschlich, Emil C.
 ; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
 ; TITLE OF INVENTION: GENES ENCODING THEM
 ; FILE REFERENCE: 040853-01-5029-03
 ; CURRENT APPLICATION NUMBER: US/10/654,528
 ; CURRENT FILING DATE: 2003-09-02
 ; PRIOR APPLICATION NUMBER: US 10/007,267
 ; PRIOR FILING DATE: 2001-12-03
 ; PRIOR APPLICATION NUMBER: US 09/333,412
 ; PRIOR FILING DATE: 1999-06-15
 ; PRIOR APPLICATION NUMBER: US 08/878,360
 ; PRIOR FILING DATE: 1997-06-18

; PRIOR APPLICATION NUMBER: US 08/683,426
 ; PRIOR FILING DATE: 1996-07-18
 ; PRIOR APPLICATION NUMBER: US 08/312,387
 ; PRIOR FILING DATE: 1994-09-24
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: Patentin version 3.2
 ; SEQ ID NO 6
 ; LENGTH: 280
 ; TYPE: PRF
 ; ORGANISM: Neisseria gonorrhoeae
 US-10-654-528-6

Query Match 69.6%; Score 1007; DB 4; Length 280;
 Best Local Similarity 74.7%; Pred. No. 1.1e-95;

Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

Qy 1 MONHVISLASAERRAHIAIDTFGRHGIPOFPDMLPSERLQAMAEIVPGLSAHPYLSG 60
 Db 1 MONHVISLASAERRAHIAIDTFGRSGIPPOFPDMLPSERLQAMAEIVPGLSAHPYLSG 60
 Qy 61 VEKACFMSHVAVLWKQALDGLPYITVFEEDVLLGGESEKFLAEDAMLOERPDPTAFIVR 120
 Db 61 VEKACFMSHVAVLWKQALDGLPYIAVFEEDVLLGGESEKFLAEDTWLERFDKDSAFIVR 120
 Qy 121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAIIISRKAMRFFLDRFALPPEGILHP 180
 Db 121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAIIISRKAMRFFLDRFALPPEGILHP 180
 Qy 181 VDLMPFSDPFDREGMPVQCLNPALCAQELHYAKFHDONSALGSLIENDRLNKKOQRDS 240
 Db 181 VDLMPFSDPFDREGMPVQSPALCTQELHYAKFSLSONSLGSDLEKD---REGRRHR 236
 Qy 241 PANTFKHRLIRALTKISRERKRRORQ 269
 Db 237 RSLKVMFDLKRALGKFKREKKRMRORQ 265

RESULT 9

US-10-472-260-176
 ; Sequence 176, Application US/10472260
 ; Publication No. US20040265328A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MICROBIOLOGICAL RESEARCH AUTHORITY
 ; APPLICANT: IMPERIAL COLLEGE INNOVATIONS LIMITED
 ; TITLE OF INVENTION: IMMUNOGENIC COMMENSAL NEISSERIA SEQUENCES
 ; FILE REFERENCE: GMS/DJC/23480
 ; CURRENT APPLICATION NUMBER: US/10/472,260
 ; CURRENT FILING DATE: 2003-09-22
 ; NUMBER OF SEQ ID NOS: 199
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 176
 ; LENGTH: 280
 ; TYPE: PRF
 ; ORGANISM: Neisseria meningitidis (group B)
 US-10-472-260-176

Query Match 64.5%; Score 933; DB 5; Length 280;
 Best Local Similarity 69.5%; Pred. No. 5.3e-88;

Matches 187; Conservative 24; Mismatches 54; Indels 4; Gaps 1;

Qy 1 MONHVISLASAERRAHIAIDTFGRHGIPOFPDMLPSERLQAMAEIVPGLSAHPYLSG 60
 Db 1 MONHVISLASAERRAHIAIDTFGVGIPPOFPDMLPSERLQAMAEIVPGLSAHPYLSG 60
 Qy 61 VEKACFMSHVAVLWKQALDGLPYITVFEEDVLLGGESEKFLAEDAMLOERPDPTAFIVR 120
 Db 61 VEKACFMSHVAVLWKQALDGLPYIAVFEEDVLLGGESEKFLAEDTWLERFDKDSAFIVR 120
 Qy 121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAIIISRKAMRFFLDRFALPPEGILHP 180
 Db 121 LETMFHVLITSPSGVADYCGRAFPILSESHWGTAIIISRKAMRFFLDRFALPPEGILHP 180
 Qy 181 VDLMPFSDPFDREGMPVQCLNPALCAQELHYAKFHDONSALGSLIENDRLNKKOQRDS 240
 Db 181 VDLMPFSDPFDREGMPVQCLNPALCAQELHYAKFHDONSALGSLIENDRLNKKOQRDS 240

Db 181 VDMWMTYFDFKGMFVYQVNPALCTQELHYAKFLSKNSWLGSDLEKD-----REQERRR 236
QY 241 PAMTFKRLRLATLTKTSREKRRQREQ 269
Db 237 RSLKVMFDLRLALGKGRKKRKRERQRQ 265

RESULT 10

US-10-398-037-1
Sequence 1, Application US/10398037
Publication No. US20040138414A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS; YUE, Henry;
APPLICANT: TANG, Y. TOM; NGUYEN, Damiel B.;
APPLICANT: VAO, Monique G.; XU, Yuming;
APPLICANT: TRIBOULET, Catherine M.; SANJANNALA, Madhusudan M.;
APPLICANT: CHAMLA, Narinder K.; BAUGHN, Mariah R.;
APPLICANT: SAPPERSTEIN, Stephanie K.; LAL, Preeti G.;
APPLICANT: THORNTON, Michael B.; GANDHI, Ameeta R.;
APPLICANT: RAMKUMAR, Jayalaxmi; ELIOTT, Vicki S.;
APPLICANT: ARVIZU, Chandra S.; THANGAVELU, Kavitha;
APPLICANT: GIEZEN, Kimberly J.; DING, Li;
APPLICANT: AU-YOUNG, Janice K.; TRAN, Bao;
APPLICANT: POLICKY, Jennifer L.; LEE, Sally;
APPLICANT: LU, Dying Aina M.; BURFORD, Neil;
APPLICANT: WARREN, Bridget A.; GURURAJAN, Rajagopal;
APPLICANT: DUGGAN, Brendan M.; HONCHELL, Cynthia D.;
APPLICANT: HAFALIA, April J.A.
TITLE OF INVENTION: SECRETED PROTEINS
FILE REFERENCE: PI-0240 USN
CURRENT APPLICATION NUMBER: US/10/398,037
PRIOR FILING DATE: 2003-03-28
PRIOR APPLICATION NUMBER: PCT/US01/30042
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/242,218
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: US 60/236,869
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: US 60/239,812
PRIOR FILING DATE: 2000-10-11
PRIOR APPLICATION NUMBER: US 60/240,108
PRIOR FILING DATE: 2000-10-12
PRIOR APPLICATION NUMBER: US 60/241,282
PRIOR FILING DATE: 2000-10-17
NUMBER OF SEQ ID NOS: 136
SOFTWARE: PERL Program
SEQ ID NO 1
LENGTH: 622
TYPE: PR
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 3211795CD1
US-10-398-037-1

Query Match 9.0%; Score 130.5; DB 4; Length 622;
Best Local Similarity 25.1%; Pred. No. 0.00029;
Matches 58; Conservative 37; Mismatches 83; Indels 53; Gaps 12;
QY 5 VLSIASAABERRAHADTFGRHGIIPQFPDL-----WPSRLBQAMALVPGLSAHFY--- 57
Db 345 MTLRRRRQDRREKMLALQAOJIECRIVEAVDGRKAMTSGVBLGIGMLFGY-RDPYHGR 403
QY 58 -LSGVEKACFMASHAVLMKQALDEGLPYITVFEDV-----LLGRGEKEFLAED 104
Db 404 PLTKGELGCELSHYNIMKKEVVDGLQKSLVFEDDLRPEIRFKRLNIMKWDVE----- 456
QY 105 AMIQERPDPTAFIVLETMFVGLTSPSGVADYCGRAP-----LLESER-WGTAGYIIS 159
Db 457 ---REGIDMDLIVYGRKMOVER-----FE-----KAVPRVNVLYEADYSYVTLAYVIS 502
QY 160 RKAMRFFLDLRFALPREGLHPVLDLMMFSDFDREGMVVCQINPALCAQELH 210

Db 503 LQARLKL---AAEPLSKLVPDVEFL-PVWFDRH--FVSEYKAHFSILRNHL 547

RESULT 11

US-10-335-977-8513
Sequence 8513, Application US/10335977
Publication No. US20040052799A1
GENERAL INFORMATION:
APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHYE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 8513:
SEQUENCE CHARACTERISTICS:
LENGTH: 332 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...332
SEQUENCE DESCRIPTION: SEQ ID NO: 8513:
US-10-335-977-8513

Query Match 8.9%; Score 128.5; DB 4; Length 332;
Best Local Similarity 23.5%; Pred. No. 0.0002; Indels 103; Gaps 17;
Matches 74; Conservative 39; Mismatches 99;
QY 29 PQFPDALMPSER-LBQAMALVPGLSA-----HPY-----LSGVEKACFMASHAVLMKQAL 77
Db 35 PQFPDALISPKHDFELQDELVNAQSLQSDWYHSVVGAGLTPLEIGCTLSHLMKECV 94
QY 78 DRLPYITVFEDVVLIGSEBEKFL-AEDAMIQERPDPTAFIVRL-----ETMPMHY 128
Db 95 KLDPQ-VVILBEDVTL---ESHFMQALBEDCLKSPD-----FVRLVGYWYHETKF-HV 144
QY 129 LNSPSVA---DYCRARFPLES----- 148
Db 145 L-PKEFVPPDPDSKKNPILKFKKFPDVSFNLSTHKVITHYILKKI QNSYATTHEK 202
QY 149 -----EHM-----GTAGYIISRKAMRFFLDLRFALPREGLHPVLDLMMFSDFPDRBGM 195
Db 203 EAPFLEHFLUTSVYVASTAGYVLTLPKGAKTFLAETESF--KILIEPDMFMDNSAHY---- 256

```
QY 136 PYCQJNPAICAOELHYAKHDONSALGSIIEHDLINRKOQRDSEFANTPEKRLPALTK 255
DB 257 DVANLTYVCPVSL-----SEHSIDSTIOKPOKSKLSYPLPQKSTFKN-----LTY 304
QY 256 ISREBKRRQREOF 270
DB 305 YSLAKKRKLMAFOY 319

RESULT 12
US-11-097-143-15177
; Sequence 15177, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Ventec, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15177
; LENGTH: 1128
; TYPE: PRF
; ORGANISM: DROSOPHILA
US-11-097-143-15177

Query Match 8.5%; Score 122.5; DB 6; Length 1128;
Best Local Similarity 25.7%; Pred. No. 0.0043;
Matches 56; Conservative 26; Mismatches 73; Indels 63; Gaps 10;

QY 5 VLSLSAARRAHIDTFGRHGIPOFPDAL---MPSRLBQAMALVPGLS---AHFY 57
DB 853 WINKRRPRRERREMERLEFIEGIEAHFPAVDGKELSTERLEMGVRFLEGGEDPYHHA 912
QY 58 LSGVEKACFMSHAVALMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERFDDTA 117
DB 913 MTMGEGGCFLSHYNIWVMWRKQKLEVLLEDDI-----RFEED----- 950
QY 118 IYRLTFMFMHVLTSFSGVD---YCGRAFPILLES-----HMGTYGII 158
DB 951 --YFQNAVRILANQANNAQYDLITFGKRLKESESPAVENADNLVHAGYSTW-TLGYI 1007
QY 159 SRKAMRFLDRFALPPEGIHPVD---LMPFSDFPDR 192
DB 1008 S---LQGALKTLAAKPLDKLIPVDFPLM-----FDR 1037

RESULT 13
US-10-723-860-1579
; Sequence 1579, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882,0193,NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1579
; LENGTH: 738
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-723-860-1579
```

```
Query Match 8.4%; Score 121; DB 5; Length 738;
Best Local Similarity 23.1%; Pred. No. 0.0035;
Matches 53; Conservative 42; Mismatches 100; Indels 34; Gaps 9;
```

```
QY 5 VLSLSAARRAHIDTFGRHGIPOFPDAL---MPSRLBQAMALVPG---LSAP 56
DB 458 WINKRRPRRERREMERLEFIEGIEAHFPAVDGKALNTSOLKALNIEMLPGRDPYSSKP 517
QY 57 YLSGVEKACFMSHAVALMKQALDEGLPYITVFEDVLLGEGEKKFLAE--DAWLQERFDD 114
DB 518 LTRG-EIGCFLSHYSVWKEVIDRELEKTLVIEDVRFHQFKKLMKMDNIDQQLDWE 576
QY 115 TAFYRLTFMFMHVLTSFSGVADYCGRAFPILLES--WGTAGYIISRKAMRFLDRFAL 173
DB 577 LIYIGRKMQVKEPPEKAVPVAN-----LVEADYSTWTLGYVLSLGAQLV---GAN 626
QY 174 PEGGLHPYDLM-----FDFPDRBGMPCQJNPAICAOELHY 211
DB 627 PRGKRLPVDFPLPVWVNHGHPVAEYKEYSRLKAFSAP-LLIYPTX 674
```

```
RESULT 14
US-10-264-237-1829
; Sequence 1829, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P431P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1829
; LENGTH: 423
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-264-237-1829

Query Match 8.2%; Score 118.5; DB 4; Length 423;
Best Local Similarity 26.2%; Pred. No. 0.003;
Matches 64; Conservative 28; Mismatches 89; Indels 63; Gaps 13;

QY 5 VLSLSAARRAHIDTFGRHGIPOFPDAL---LMPSRLEQAMALVPGSAHPY--- 57
DB 155 VLSLARPRRERREMERLEFIEGIEAHFPAVDGKALNTSOLKALNIEMLPGRDPYSSKP 213
QY 58 -LSGVEKACFMSHAVALMKQALDEGLPYITVFEDVLLGEGEKKFLAEDAWLQERFDDTA 116
DB 214 TLTKGEVCGFLSHYSIMBEVVARGLARVLVDFDV-----RFSN-- 253
```

QY 117 FIVRLTETMFHVLTSF-SGVADYCGRAFPILSE-----HMGTAGYIISR 160
 DB 254 FRGRLERLMEVDYAEKLSMDLIYIGRKQVPEKETAVEGLPGLVVAGYSYV-TLAYALRL 312
 QY 161 KAMRFFLDRFALPPEGLHPVD---LMF-----SDFPDREGMPVCQINPALCAOE 208
 DB 313 AGARKLL---ASQPLRRMLPVDLPIFMDQHPNEQYKAHFWPRD-LVAFSAQPLLA--P 367
 QY 209 LHYA 212
 DB 368 THYA 371

RESULT 15
 US-10-723-860-4111
 ; Sequence 4111, Application US/10723860
 ; Publication NO. US20040253606A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsburg, Wendy M.
 ; APPLICANT: Zlotnick, Albert
 ; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
 ; FILE REFERENCE: 05862.0193.NPUS01
 ; CURRENT FILING DATE: 2003-11-26
 ; PRIOR APPLICATION NUMBER: 60/429,739
 ; PRIOR FILING DATE: 2002-11-26
 ; NUMBER OF SEQ ID NOS: 8393
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 4111
 ; LENGTH: 560
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-723-860-4111

Query Match 8.2%; Score 118.5; DB 5; Length 560;
 Best Local Similarity 26.2%; Pred. No. 0.0044;
 Matches 64; Conservative 28; Mismatches 89; Indels 63; Gaps 13;

QY 5 VISLSAARRAHIADTFGRHGIPOFPDA---IMPSERLQAMALVPGLSAHFY--- 57
 DB 323 VISLARPPRRRRBSMLASLWEMEISGRVVDVAVDGMNINSSAIRVIGVDLPGYQ-DPYSGR 381
 QY 58 -ISGVKACGFMGSHAVLWKQALDEGLPYITVFEDDVLGSGEKEKTLAEDAMLOERPDPTA 116
 DB 382 TIKGIVGCFLSHSYISWEEVARGLARVLVFEEDV-----RFRSN-- 421
 QY 117 FIVRLTETMFHVLTSF-SGVADYCGRAFPILSE-----HMGTAGYIISR 160
 DB 422 FRGRLERLMEVDYAEKLSMDLIYIGRKQVPEKETAVEGLPGLVVAGYSYV-TLAYALRL 480
 QY 161 KAMRFFLDRFALPPEGLHPVD---LMF-----SDFPDREGMPVCQINPALCAOE 208
 DB 481 AGARKLL---ASQPLRRMLPVDLPIFMDQHPNEQYKAHFWPRD-LVAFSAQPLLA--P 535
 QY 209 LHYA 212
 DB 536 THYA 539

Search completed: April 7, 2006, 15:45:08
 Job time : 170 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:42:34 ; Search time 25 Seconds

(without alignments)
343.112 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446

Sequence: 1 MONHYSLASAAERRAHAD.....ISRRKKRQRREQIVPRQ 275

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 184161 seqs, 31191982 residues

Total number of hits satisfying chosen parameters: 184161

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSS/ptodata/1/pubpaa/US08 NEW PUB.pep:*
2: /SIDSS/ptodata/1/pubpaa/US06 NEW PUB.pep:*
3: /SIDSS/ptodata/1/pubpaa/US07 NEW PUB.pep:*
4: /SIDSS/ptodata/1/pubpaa/PCT_NEW PUB.pep:*
5: /SIDSS/ptodata/1/pubpaa/US09 NEW PUB.pep:*
6: /SIDSS/ptodata/1/pubpaa/US10 NEW PUB.pep:*
7: /SIDSS/ptodata/1/pubpaa/US11 NEW PUB.pep:*
8: /SIDSS/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1288	89.1	279	7	US-11-102-497-8
2	1281	88.6	279	6	US-10-467-657-1310
3	1007	69.6	280	7	US-11-102-497-6
4	996	68.9	280	6	US-10-467-657-1336
5	118.5	8.2	327	6	US-10-821-234-884
6	98	6.8	481	7	US-11-827-177-1
7	88.5	6.1	651	7	US-11-087-099-2611
8	87.5	6.1	238	7	US-11-239-510-21
9	87.5	6.1	290	7	US-11-032-773-957
10	87.5	6.1	319	7	US-11-032-773-955
11	84	5.8	395	7	US-11-084-220-2
12	83.5	5.8	480	7	US-11-239-510-11
13	83.5	5.8	443	7	US-11-239-510-11
14	83.5	5.8	353	6	US-10-506-454-1626
15	82	5.7	120	6	US-10-932-334-77
16	81.5	5.6	416	6	US-10-501-035-279
17	80.5	5.6	571	7	US-11-087-099-7090
18	79.5	5.5	119	7	US-11-226-886-20
19	79.5	5.5	119	7	US-11-159-046-4
20	79.5	5.5	248	6	US-10-512-184-32
21	79.5	5.5	317	6	US-10-512-184-69
22	79	5.5	425	7	US-11-096-568A-20512
23	79	5.5	645	7	US-11-096-568A-19117
24	79	5.5	648	7	US-11-096-568A-19116
25	79	5.5	711	7	US-11-096-568A-19115

26	78	5.4	331	7	US-11-087-099-3050	Sequence 3050, Ap
27	77.5	5.4	968	6	US-10-501-035-219	Sequence 219, App
28	77	5.3	363	7	US-11-121-731A-1	Sequence 1, Appl
29	77	5.3	539	6	US-10-793-626-340	Sequence 340, App
30	77	5.3	805	6	US-10-927-641-77	Sequence 77, Appl
31	76	5.3	118	6	US-10-932-334-75	Sequence 75, Appl
32	75.5	5.2	434	7	US-11-087-099-11503	Sequence 11503, A
33	75.5	5.2	687	6	US-10-485-517-274	Sequence 274, App
34	75	5.2	252	7	US-11-072-512-2120	Sequence 2120, App
35	75	5.2	268	7	US-11-096-568A-2036	Sequence 2036, Ap
36	75	5.2	805	6	US-10-485-517-188	Sequence 198, App
37	75	5.2	842	7	US-11-096-568A-34417	Sequence 34417, A
38	75	5.2	869	7	US-11-096-568A-34416	Sequence 34416, A
39	75	5.2	970	7	US-10-932-334-79	Sequence 34415, A
40	74	5.1	120	6	US-10-932-334-79	Sequence 79, Appl
41	74	5.1	303	7	US-11-096-568A-14763	Sequence 14763, A
42	74	5.1	395	7	US-11-096-568A-14762	Sequence 14762, A
43	74	5.1	411	7	US-11-096-568A-14761	Sequence 14761, A
44	73.5	5.1	302	7	US-11-165-226-137	Sequence 127, App
45	73.5	5.1	436	7	US-11-096-568A-21599	Sequence 21599, A

ALIGNMENTS

RESULT 1
US-11-102-497-8
; Sequence 8, Application US/11102497
; Publication No. US20050271690A1
; GENERAL INFORMATION:
; APPLICANT: Gotschlich, Emil C.
; TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
; FILE REFERENCE: 040853-01-5029-03
; CURRENT FILING DATE: 2005-04-08
; PRIOR FILING DATE: 2003-09-02
; PRIOR FILING DATE: 2003-09-02
; PRIOR FILING DATE: 2001-12-03
; PRIOR FILING DATE: 2001-12-03
; PRIOR FILING DATE: 1999-06-15
; PRIOR FILING DATE: 1999-06-15
; PRIOR FILING DATE: 1997-06-18
; PRIOR FILING DATE: 1997-06-18
; PRIOR FILING DATE: 1996-07-18
; PRIOR FILING DATE: 1996-07-18
; PRIOR FILING DATE: 1994-09-24
; PRIOR FILING DATE: 1994-09-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-11-102-497-8
Query Match 89.1%; Score 1288; DB 7; Length 279;
Best Local Similarity 90.0%; Pred. No. 5.4e-120;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;
QY 1 MONHYSLASAAERRAHADTFGRGRTGFPOFDLMPSERLBOMAEIVPGLSAHPIYSG 60
DB 1 MONHYSLASAAERRAHADTFGRGRTGFPOFDLMPSERLBOMAEIVPGLSAHPIYSG 60
QY 61 VEKACFMSHAYLWKOALDEGLPYITVPEBDVLGSGEKEFLAEDAMLOERFPDPTAFIVR 120
DB 61 VEKACFMSHAYLWKOALDEGLPYITVPEBDVLGSGEKEFLAEDAMLOERFPDPTAFIVR 120
QY 121 LETMFMHTLTPSGVADYCGRAFLPBBSEHNGTGYITSRKAMFFLDRLPALPPEGALP 180
DB 121 LETMFMHTLTPSGVADYCGRAFLPBBSEHNGTGYITSRKAMFFLDRLPALPPEGALP 180
QY 181 VDLMMFSDFFPREGMPVCQALPALCAQELHYAKFHQDSALGSLIEHRLNLRKQQRDS 240

Db 161 VDLMMFGNDDREGMPVCOINPALCAQELHYAKFHDONSALSLIEHDLRLNRKQWRDS 240

Qy 241 PANTFPHRLIRALTITKISRERKRQRREQ----FTVFPQ 275

Db 241 PANTFPHRLIRALTITKISRERKRQRREQGLIKIIVFPQ 279

RESULT 2

US-10-467-657-1310

Sequence 1310, Application US/10467657

Publication No. US20050260581A1

GENERAL INFORMATION:

APPLICANT: CHIRON SPA

APPLICANT: FONTANA Maria Rita

APPLICANT: PIZZA Mariagrazia

APPLICANT: MASTIGNANI Vega

APPLICANT: MONACI Elisabetta

TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS

FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/10/467,657

CURRENT FILING DATE: 2003-08-11

PRIOR APPLICATION NUMBER: GB-0103424.8

PRIOR FILING DATE: 2001-02-12

NUMBER OF SEQ ID NOS: 9218

SOFTWARE: SeqWin99, version 1.04

SEQ ID NO 1310

LENGTH: 279

TYPE: PRT

ORGANISM: Neisseria gonorrhoeae

US-10-467-657-1310

Query Match 88.6%; Score 1281; DB 6; Length 279;

Best Local Similarity 90.0%; Pred. No. 2,7e-119;

Matches 251; Conservative 6; Mismatches 18; Indels 4; Gaps 1;

Qy 1 MONHVLSLSAARRAHIAIDTFGRHGIPPOFPDALMPSERLBOAMAEIVPGISAHPIYLSG 60

Db 1 MONHVLSLSAARRAHIAIDTFGRHGIPPOFPDALMPSERLBOAMAEIVPGISAHPIYLSG 60

Qy 61 VERACFMSHAIVLWKQALDEGLPYITVPEDDVLLGEGEKKFLAEDAWLQERPDPTAFIVR 120

Db 61 VERACFMSHAIVLWKQALDEGLPYITVPEDDVLLGEGEKKFLAEDAWLQERPDPTAFIVR 120

Qy 121 LETMFPHVLTSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFFLDRFAALPPEGILHP 180

Db 121 LETMFPHVLTSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFFLDRFAALPPEGILHP 180

Qy 181 VDLMMFSDPFDEGMPVCOINPALCAQELHYAKFHDONSALSLIEHDLRLNRKQWRDS 240

Db 181 VDLMMFSDPFDEGMPVCOINPALCAQELHYAKFHDONSALSLIEHDLRLNRKQWRDS 240

Qy 241 PANTFPHRLIRALTITKISRERKRQRREQ----FTVFPQ 275

Db 241 PANTFPHRLIRALTITKISRERKRQRREQITGKIIVFPQ 279

RESULT 3

US-11-102-497-6

Sequence 6, Application US/11102497

Publication No. US20050271690A1

GENERAL INFORMATION:

APPLICANT: Gotschlich, Emil C.

TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND

FILE REFERENCE: 040853-01-5029-03

CURRENT APPLICATION NUMBER: US/11/102,497

CURRENT FILING DATE: 2005-04-08

PRIOR APPLICATION NUMBER: US/10/654,528

PRIOR FILING DATE: 2003-09-02

PRIOR APPLICATION NUMBER: US 10/007,267

PRIOR FILING DATE: 2001-12-03

PRIOR APPLICATION NUMBER: US 09/333,412

PRIOR FILING DATE: 1999-06-15

PRIOR APPLICATION NUMBER: US 08/878,360

PRIOR FILING DATE: 1997-06-18

PRIOR APPLICATION NUMBER: US 08/683,426

PRIOR FILING DATE: 1996-07-18

PRIOR APPLICATION NUMBER: US 08/312,387

PRIOR FILING DATE: 1994-09-24

NUMBER OF SEQ ID NOS: 13

SOFTWARE: PatentIn version 3.2

SEQ ID NO 6

LENGTH: 280

TYPE: PRT

ORGANISM: Neisseria gonorrhoeae

US-11-102-497-6

Query Match 69.6%; Score 1007; DB 7; Length 280;

Best Local Similarity 74.7%; Pred. No. 3.8e-92;

Matches 201; Conservative 20; Mismatches 44; Indels 4; Gaps 1;

Qy 1 MONHVLSLSAARRAHIAIDTFGRHGIPPOFPDALMPSERLBOAMAEIVPGISAHPIYLSG 60

Db 1 MONHVLSLSAARRAHIAIDTFGRHGIPPOFPDALMPSERLBOAMAEIVPGISAHPIYLSG 60

Qy 61 VERACFMSHAIVLWKQALDEGLPYITVPEDDVLLGEGEKKFLAEDAWLQERPDPTAFIVR 120

Db 61 VERACFMSHAIVLWKQALDEGLPYITVPEDDVLLGEGEKKFLAEDAWLQERPDPTAFIVR 120

Qy 121 LETMFPHVLTSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFFLDRFAALPPEGILHP 180

Db 121 LETMFPHVLTSPSGVADYCGRAFPILSESHMGTAGYIISRKAMRFFLDRFAALPPEGILHP 180

Qy 181 VDLMMFSDPFDEGMPVCOINPALCAQELHYAKFHDONSALSLIEHDLRLNRKQWRDS 240

Db 181 VDLMMFSDPFDEGMPVCOINPALCAQELHYAKFHDONSALSLIEHDLRLNRKQWRDS 240

Qy 241 PANTFPHRLIRALTITKISRERKRQRREQ 269

Db 237 RSLKVMFDLKRALGKFGREKKRQRREQ 265

RESULT 4

US-10-467-657-1326

Sequence 1326, Application US/10467657

Publication No. US20050260581A1

GENERAL INFORMATION:

APPLICANT: CHIRON SPA

APPLICANT: FONTANA Maria Rita

APPLICANT: PIZZA Mariagrazia

APPLICANT: MASTIGNANI Vega

APPLICANT: MONACI Elisabetta

TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS

FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/10/467,657

CURRENT FILING DATE: 2003-08-11

PRIOR APPLICATION NUMBER: GB-0103424.8

PRIOR FILING DATE: 2001-02-12

NUMBER OF SEQ ID NOS: 9218

SOFTWARE: SeqWin99, version 1.04

SEQ ID NO 1326

LENGTH: 280

TYPE: PRT

ORGANISM: Neisseria gonorrhoeae

US-10-467-657-1326

Query Match 68.9%; Score 996; DB 6; Length 280;

Best Local Similarity 74.3%; Pred. No. 4.7e-91;

Matches 200; Conservative 19; Mismatches 46; Indels 4; Gaps 1;

Qy 1 MONHVLSLSAARRAHIAIDTFGRHGIPPOFPDALMPSERLBOAMAEIVPGISAHPIYLSG 60

Db 1 MONHVLSLSAARRAHIAIDTFGRHGIPPOFPDALMPSERLBOAMAEIVPGISAHPIYLSG 60

Qy 61 VERACFMSHAIVLWKQALDEGLPYITVPEDDVLLGEGEKKFLAEDAWLQERPDPTAFIVR 120

```

Db      61 VERACFSHVAWLMKQALDEGLPYIAVEDDVLGKQAEKFLAEDTWLEERFDKOSATVR 120
Qy      121 LETMFHVLTPSPGVADYCCRAAPLLESEMGTGIIISKAMFFLDRAALPREGLP 180
      121 LETMFARVIRPDVILNVEENRSPFLSESCGTAGYIISBAMFFLDRAVLPERIK 180
Qy      181 VDLMPFSDPDRGMPVCOINPALCAQELHYAKPHDONSAIGLIEHDLINRKOQRDS 240
      181 VDLMPFTYFPDKGMPVYQVSPALCTQELHYAKFLSQNSMLGSDLEKD---REGRRHR 236
Qy      241 PANTFKRLIRALTKISREERKORREQ 269
      237 RSLKVMFDLRALGKFGREKKERORQ 265

```

RESULT 5
US-10-821-234-884

```

; Sequence 884, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarman, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 884
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-884

```

Query Match
Best Local Similarity 26.2%; Score 118.5; DB 6; Length 327;
Pred. No. 0.00052; Matches 64; Conservative 28; Mismatches 89; Indels 63; Gaps 13;

```

Qy      5 VISIAGAERRAHIAITFGRHGIPQFPDA---LMPSERLBQAMLEIVPGLSAHNY--- 57
      40 VISIARPRRRKEMLASISWMEISGRVDAVDGMMNLSSAIRNLGVDLRPGIQ-DPYSGR 98
Qy      58 -LSGVKACFMSHVAWLMKQALDEGLPYITVFEDDVLGEGEKEFLAEDAMLOERFPDPA 116
      99 TLTKGVGCFLSHYSIWEEVVAAGLARVLVFEDD-----RFESEN-- 138
Qy      117 FIVLEETMTHVLTSP-SGVADYCCRAAPLLESE-----HWGTAGYIISR 160
      139 FGRLEERLMEDEVAEKLISWDLITLGRKQVPEKEITAVEGLPGLVVAGYSYM-TLAVYLR 197
Qy      161 KAMRFIDRPAALPREGLPVD---LMMF-----SDPDRGMPVCOINPALCAQ 208
      198 AGARGL---ASQPLRRMLPVDLFLPIMDQHPNEQYKAFWPRD-LVAPSAQPLIAA-P 252
Qy      209 LHYA 212
      253 THYA 256
Db

```

RESULT 6
US-11-227-177-1
; Sequence 1, Application US/11227177
; Publication No. US20060014262A1
; GENERAL INFORMATION:
; APPLICANT: KYOMA HAKKO KOGYO CO., LTD.
; TITLE OF INVENTION: Beta 1,3-galactosyltransferase and a DNA coding for said enzyme
; FILE REFERENCE: 11428WO1
; CURRENT APPLICATION NUMBER: US/11/227,177

```

; CURRENT FILING DATE: 2005-09-16
; PRIOR APPLICATION NUMBER: US/10/475,587
; PRIOR FILING DATE: 2003-10-23
; PRIOR APPLICATION NUMBER: JP 2001-123864
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-11-227-177-1

```

Query Match
Best Local Similarity 30.2%; Score 98; DB 7; Length 481;
Pred. No. 0.0092; Matches 32; Conservative 16; Mismatches 36; Indels 22; Gaps 6;

```

Qy      29 FQFPDALMPS-ERLEQAMAEIVGSL-----AHPYISGVKACFMSHVAWLMKQ 76
      35 FQIFDAISPKHDFEKLQELVDSSNLKSDWPHSDYCYQELLPQFPGCTLSHYLMKEC 94
Db      95 VKTNQP-VVILEDDIAL---ESNFMQALBDCIKSPD-----FVRL 131
Qy      77 LDEGLPYITVFEDDVLGEGEKEFL-AEDAMLOERFPDPAFVRL 121
      95 VKTNQP-VVILEDDIAL---ESNFMQALBDCIKSPD-----FVRL 131
Db

```

RESULT 7
US-11-087-099-2611

```

; Sequence 2611, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2611
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Brucella melitensis bioovar Abortus
US-11-087-099-2611

```

Query Match
Best Local Similarity 24.9%; Score 88.5; DB 7; Length 651;
Pred. No. 1.2; Matches 50; Conservative 23; Mismatches 65; Indels 63; Gaps 11;

```

Qy      89 DVLGEGEKEFLAEDAMLOERFPDPAFVRLLETMFHVLTPSPGVADYCCRAF---PL 145
      299 DGIIGDSEFPGAERREIED-YQPD-----ELEKX-----RPAASDVLGRAMEKDP 347
Db      146 L-----ESEH---WGTAGYIISRKAMFFLDREPALP-----PQGLHFDLM 184
      348 IIVMGEDVIRFAGVSGF--TNALLEFPDRVLAMPVLENGFTVALGAALGLRVVBI 405
Qy      185 MNSDF-----FDRGMPVCOINPALCAQELHYAKPHDON-SALGSLIE 226
      406 MEGDFCFVAADQIANGISKVRMFGDPVPIVMEVRSVPHYGSGSHGSDSALFGMP 465
Qy      227 HDRLNRRKQQRDSPANTPKH 247
      466 GMRVY-----SPTNAFDY 478
Db

```

RESULT 8
US-11-239-510-21
; Sequence 21, Application US/11239510
; Publication No. US20060063715A1
; GENERAL INFORMATION:
; APPLICANT: Whitlow, Marc
; APPLICANT: Wood, James F.
; APPLICANT: Hardman, Karl
; APPLICANT: Bird, Robert

DB 103 YYYDNDVGDYWG 114

RESULT 11

US-11-084-220-2
Sequence 2, Application US/11084220
Publication No. US20050266439A1
GENERAL INFORMATION:
APPLICANT: Hjortleifsdottir, Sigridur
APPLICANT: Thordisdottir, Audur
APPLICANT: Aevartsson, Arnthor
APPLICANT: Hreysvoldsson, Gudmundur O.
APPLICANT: Blondal, Thorarinn
TITLE OF INVENTION: Thermostable Ligase from Thermus Phage
FILE REFERENCE: 2138.001US1
CURRENT APPLICATION NUMBER: US/11/084,220
PRIOR FILING DATE: 2005-03-18
PRIOR APPLICATION NUMBER: US 60/412,461
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 395
TYPE: PRF
ORGANISM: Bacteriophage TS2126
US-11-084-220-2

Query Match 5.8%; Score 84; DB 7; Length 395;
Best Local Similarity 25.6%; Pred. No. 1.7;
Matches 52; Conservative 22; Mismatches 63; Indels 66; Gaps 12;

QY 29 FQFFDLMPSEKLEQ-----MAELVPGLSAHPYLSGVKACFMSHAV-----L 72
DB 85 FNGEELAGEBEAFKAFRDSWVPLFVAEKVDGYLAQVLDGGELRPSRSISLMPPLVGL 144
QY 73 WKQALDEGLPYIVFEDDVLLGEGEKKFLAEDA--W--LQERPDPTAFIVLETFMFIY 128
DB 145 LKRAYE-----BAMARLG---KLAAEGGRTALLLEVDPAPVWVPYQERGVL 192
QY 129 LNSPS-GVADYC--GRAFPLLS--EHW-----GTAGYIIS----- 159
DB 193 LALRSIGEGHYLLPGHFFPLPEALRYVRWEPMDPDPHFRGRIRDLQGVGVVVDGAE 252
QY 160 ----RKARPFLLDRFAALPREG 178
DB 253 FVFKTGMAFLARF-LMDPEGV 274

RESULT 12

US-11-239-510-11
Sequence 11, Application US/11239510
Publication No. US20060063715A1
GENERAL INFORMATION:
APPLICANT: Whitlow, Marc
APPLICANT: Wood, James P.
APPLICANT: Hardman, Karl
APPLICANT: Bird, Robert
APPLICANT: Filpula, David
APPLICANT: Rollence, Michelle
TITLE OF INVENTION: Multivalent Antigen-Binding Proteins
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, NW
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/239,510
FILING DATE: 29-Sep-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/137,297
FILING DATE: 03-May-2002
APPLICATION NUMBER: US/09/443,213
FILING DATE: <Unknown>
APPLICATION NUMBER: US 09/166,094
FILING DATE: 05-OCT-1998
APPLICATION NUMBER: US 08/392,338
FILING DATE: 22-FEB-1995
APPLICATION NUMBER: US 07/989,846
FILING DATE: 20-NOV-1992
APPLICATION NUMBER: US 07/796,936
FILING DATE: 25-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Goldstein, Jorge A.
REGISTRATION NUMBER: 29,021
REFERENCE/DOCKET NUMBER: 0977.003000E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-11-239-510-11

Query Match 5.8%; Score 83.5; DB 7; Length 240;
Best Local Similarity 29.8%; Pred. No. 1;
Matches 37; Conservative 19; Mismatches 37; Indels 31; Gaps 9;

QY 40 RLEQMAHEIV-PGLSAHPYLSGVKAC-----FMSHAVLW-KQALDEGLPYIVT---E 88
DB 129 QIQQSDAEIVKPGAS-----VKISCRASGYTFDHAHWKQNDGGLWIGYSPGN 181
QY 89 DVLVAGEGEKKFLAEDAWLQERPDPTAFIVLETFMFIYLSPSGVADYCGRAFPPLLS 148
DB 182 DPF---KTYERPKGK-ATLTADKSSSTAYVQ-----LNSLISEDSAYVFTCRS---LNM 228
QY 149 EHWG 152
DB 229 AYWG 232

RESULT 13

US-11-239-510-19
Sequence 19, Application US/11239510
Publication No. US20060063715A1
GENERAL INFORMATION:
APPLICANT: Whitlow, Marc
APPLICANT: Wood, James P.
APPLICANT: Hardman, Karl
APPLICANT: Bird, Robert
APPLICANT: Filpula, David
APPLICANT: Rollence, Michelle
TITLE OF INVENTION: Multivalent Antigen-Binding Proteins
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, NW
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

Mon Apr 10 08:45:28 2006

us-09-211-691-2.rapbn

Page 7

Db 105 YDGDYWG 111

Search completed: April 7, 2006, 15:45:38
Job time : 26 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:30:46 ; Search time 573 Seconds

(without alignments)
663.242 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446
Sequence: 1 MONHVISLASAARRAHAD.....ISREKRRQRGRQIVPRQ 275

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 7861189 seqs, 1381955077 residues

Total number of hits satisfying chosen parameters: 7861189

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/US066_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US066_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US073_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US073_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US073_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US073_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US073_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US073_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US073_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US073_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
33: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
34: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
35: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
36: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
37: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
38: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
39: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
40: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
41: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
42: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
43: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*

44: /cgn2_6/ptodata/1/paa/US600_COMB.pep.*
45: /cgn2_6/ptodata/1/paa/US601_COMB.pep.*
46: /cgn2_6/ptodata/1/paa/US602_COMB.pep.*
47: /cgn2_6/ptodata/1/paa/US603_COMB.pep.*
48: /cgn2_6/ptodata/1/paa/US604_COMB.pep.*
49: /cgn2_6/ptodata/1/paa/US605_COMB.pep.*
50: /cgn2_6/ptodata/1/paa/US606_COMB.pep.*
51: /cgn2_6/ptodata/1/paa/US607_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1446	100.0	275	22	US-09-211-691-2	Sequence 2, Appl1
2	1446	100.0	275	33	US-10-317-428-2	Sequence 2, Appl1
3	1446	100.0	275	33	US-10-317-773-2	Sequence 2, Appl1
4	1288	89.1	279	36	US-10-654-528-8	Sequence 8, Appl1
5	1288	89.1	279	41	US-11-102-497-8	Sequence 8, Appl1
6	1281	88.6	279	34	US-10-467-657-1310	Sequence 1310, Ap
7	1281	88.6	279	34	US-10-467-657A-1310	Sequence 1310, Ap
8	1007	69.6	280	30	US-10-096-129-6	Sequence 6, Appl1
9	1007	69.6	280	36	US-10-654-528-6	Sequence 6, Appl1
10	1007	69.6	280	41	US-11-102-497-6	Sequence 6, Appl1
11	996	68.9	280	34	US-10-467-657-1326	Sequence 1326, Ap
12	996	68.9	280	34	US-10-467-657A-1326	Sequence 1326, Ap
13	933	64.5	312	34	US-10-472-260-116	Sequence 176, App
14	436	30.2	312	50	US-60-691-214-3189	Sequence 3189, App
15	354	25.2	255	36	US-10-603-108-2407	Sequence 2407, App
16	364	25.2	255	45	US-60-128-476-4128	Sequence 4128, App
17	300	20.7	282	50	US-60-691-214-3405	Sequence 3405, App
18	255	17.6	50	20	US-09-084-616-2	Sequence 2, Appl1
19	239	16.5	50	20	US-09-084-616-3	Sequence 3, Appl1
20	130.5	9.0	278	39	US-10-912-745-281	Sequence 281, App
21	130.5	9.0	278	39	US-10-912-745A-281	Sequence 281, App
22	130.5	9.0	278	49	US-60-512-690-418	Sequence 418, App
23	130.5	9.0	278	49	US-60-519-832-418	Sequence 418, App
24	130.5	9.0	278	49	US-60-552-390-418	Sequence 418, App
25	130.5	9.0	288	27	US-09-758-446-1206	Sequence 1206, App
26	130.5	9.0	288	32	US-10-212-759-1206	Sequence 1206, App
27	130.5	9.0	466	31	US-10-170-205E-36168	Sequence 36168, A
28	130.5	9.0	521	32	US-10-218-140-3328	Sequence 3328, App
29	130.5	9.0	622	26	US-09-611-523-372	Sequence 372, App
30	130.5	9.0	622	26	US-09-611-526-3054	Sequence 3054, App
31	130.5	9.0	622	31	US-10-170-205E-36147	Sequence 36147, A
32	130.5	9.0	622	33	US-10-305-278-272	Sequence 272, App
33	130.5	9.0	622	33	US-10-398-037-1	Sequence 1, Appl1
34	130.5	9.0	622	39	US-10-912-745-283	Sequence 283, App
35	130.5	9.0	622	39	US-10-912-745A-283	Sequence 283, App
36	130.5	9.0	622	46	US-60-236-869-1	Sequence 1, Appl1
37	130.5	9.0	622	48	US-60-453-050-9797	Sequence 9797, App
38	130.5	9.0	622	48	US-60-453-135-9797	Sequence 9797, App
39	130.5	9.0	622	48	US-60-466-412-9797	Sequence 9797, App
40	130.5	9.0	622	49	US-60-512-690-420	Sequence 420, App
41	130.5	9.0	622	49	US-60-519-832-420	Sequence 420, App
42	130.5	9.0	645	39	US-60-552-390-420	Sequence 420, App
43	130.5	9.0	645	39	US-10-912-745-280	Sequence 280, App
44	130.5	9.0	645	39	US-10-912-745A-280	Sequence 280, App
45	130.5	9.0	645	49	US-60-512-690-417	Sequence 417, App

ALIGNMENTS

RESULT 1
US-09-211-691-2
; Sequence 2, Application US/09211691
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel

APPLICANT: Young, N. Martin
APPLICANT: Makarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Fusion Proteins for Use in Enzymatic Synthesis of
FILE REFERENCE: 01957-012910US
CURRENT APPLICATION NUMBER: US/09/211,691
CURRENT FILING DATE: 1998-12-14
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 275
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-09-211-691-2

Query Match 100.0%; Score 1446; DB 22; Length 275;
Best Local Similarity 100.0%; Pred. No. 4e-137;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MONHVISLASAARRAHADTFGRHGIPPOFPDAMPSERLQAMAEIVPGLSAHPYLSG 60
DB 1 MONHVISLASAARRAHADTFGRHGIPPOFPDAMPSERLQAMAEIVPGLSAHPYLSG 60
QY 61 VEKACFMSHAVALMKQALDEGLPYITVFEDVDLLGSEGEKFLAEDAWLOERPPDTAFIVR 120
DB 61 VEKACFMSHAVALMKQALDEGLPYITVFEDVDLLGSEGEKFLAEDAWLOERPPDTAFIVR 120
QY 121 LETMFMAHVLTPSPGVADYCGRAFPILLESHEHWGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMAHVLTPSPGVADYCGRAFPILLESHEHWGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
QY 181 VDLMPFSDPFDEBEGPVCQALNPALCAOELHYAKFHDONSALGSLIEHDLRLNRKQORRDS 240
DB 181 VDLMPFSDPFDEBEGPVCQALNPALCAOELHYAKFHDONSALGSLIEHDLRLNRKQORRDS 240
QY 241 PANTFKHRLIRALTISRERKRQRROEQFIVPFQ 275
DB 241 PANTFKHRLIRALTISRERKRQRROEQFIVPFQ 275

RESULT 2
US-10-317-428-2
Query Match 100.0%; Score 1446; DB 33; Length 275;
Best Local Similarity 100.0%; Pred. No. 4e-137;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Makarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
FILE REFERENCE: 019633-000811US
CURRENT APPLICATION NUMBER: US/10/317,428
CURRENT FILING DATE: 2002-12-11
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
PRIOR APPLICATION NUMBER: US 09/211,691
PRIOR FILING DATE: 1998-12-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 275
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-10-317-428-2

QY 1 MONHVISLASAARRAHADTFGRHGIPPOFPDAMPSERLQAMAEIVPGLSAHPYLSG 60
DB 1 MONHVISLASAARRAHADTFGRHGIPPOFPDAMPSERLQAMAEIVPGLSAHPYLSG 60

DB 1 MONHVISLASAARRAHADTFGRHGIPPOFPDAMPSERLQAMAEIVPGLSAHPYLSG 60
QY 61 VEKACFMSHAVALMKQALDEGLPYITVFEDVDLLGSEGEKFLAEDAWLOERPPDTAFIVR 120
DB 61 VEKACFMSHAVALMKQALDEGLPYITVFEDVDLLGSEGEKFLAEDAWLOERPPDTAFIVR 120
QY 121 LETMFMAHVLTPSPGVADYCGRAFPILLESHEHWGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMAHVLTPSPGVADYCGRAFPILLESHEHWGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
QY 181 VDLMPFSDPFDEBEGPVCQALNPALCAOELHYAKFHDONSALGSLIEHDLRLNRKQORRDS 240
DB 181 VDLMPFSDPFDEBEGPVCQALNPALCAOELHYAKFHDONSALGSLIEHDLRLNRKQORRDS 240
QY 241 PANTFKHRLIRALTISRERKRQRROEQFIVPFQ 275
DB 241 PANTFKHRLIRALTISRERKRQRROEQFIVPFQ 275

RESULT 3
US-10-317-773-2
Query Match 100.0%; Score 1446; DB 33; Length 275;
Best Local Similarity 100.0%; Pred. No. 4e-137;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Makarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
FILE REFERENCE: 019633-000812US
CURRENT APPLICATION NUMBER: US/10/317,773
CURRENT FILING DATE: 2002-12-11
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
PRIOR APPLICATION NUMBER: US 09/211,691
PRIOR FILING DATE: 1998-12-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 275
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-10-317-773-2

QY 1 MONHVISLASAARRAHADTFGRHGIPPOFPDAMPSERLQAMAEIVPGLSAHPYLSG 60
DB 1 MONHVISLASAARRAHADTFGRHGIPPOFPDAMPSERLQAMAEIVPGLSAHPYLSG 60
QY 61 VEKACFMSHAVALMKQALDEGLPYITVFEDVDLLGSEGEKFLAEDAWLOERPPDTAFIVR 120
DB 61 VEKACFMSHAVALMKQALDEGLPYITVFEDVDLLGSEGEKFLAEDAWLOERPPDTAFIVR 120
QY 121 LETMFMAHVLTPSPGVADYCGRAFPILLESHEHWGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
DB 121 LETMFMAHVLTPSPGVADYCGRAFPILLESHEHWGTAGYIISRKAMRFFLDRFALPPEGGLHP 180
QY 181 VDLMPFSDPFDEBEGPVCQALNPALCAOELHYAKFHDONSALGSLIEHDLRLNRKQORRDS 240
DB 181 VDLMPFSDPFDEBEGPVCQALNPALCAOELHYAKFHDONSALGSLIEHDLRLNRKQORRDS 240
QY 241 PANTFKHRLIRALTISRERKRQRROEQFIVPFQ 275
DB 241 PANTFKHRLIRALTISRERKRQRROEQFIVPFQ 275

RESULT 4
US-10-654-528-8
Query Match 100.0%; Score 1446; DB 33; Length 275;
Best Local Similarity 100.0%; Pred. No. 4e-137;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Makarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
FILE REFERENCE: 019633-000813US
CURRENT APPLICATION NUMBER: US/10/654,528
CURRENT FILING DATE: 2002-12-11
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
PRIOR APPLICATION NUMBER: US 09/211,691
PRIOR FILING DATE: 1998-12-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 275
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-10-654-528-8

APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
TITLE OF INVENTION: GENES ENCODING THEM
FILE REFERENCE: 040853-01-5029-03
CURRENT APPLICATION NUMBER: US/10/654,528
CURRENT FILING DATE: 2003-09-02
PRIOR APPLICATION NUMBER: US 10/007,267
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: US 09/333,412
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: US 08/878,360
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: US 08/683,426
PRIOR FILING DATE: 1996-07-18
PRIOR APPLICATION NUMBER: US 08/312,387
PRIOR FILING DATE: 1994-09-24
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.2
SEQ ID NO 8
LENGTH: 279
TYPE: PR1
ORGANISM: Neisseria gonorrhoeae
US-10-654-528-8

Query Match 89.1%; Score 1288; DB 36; Length 279;
Best Local Similarity 90.0%; Pred. No. 4,2e-121;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY 1 MONHVLSLSAARRRAHIAIDTFRHGIPPOFPDAMPSERLQAMAEVPGLSAHPYLSG 60
DB 1 MONHVLSLSAARRRAHIAIDTFRHGIPPOFPDAMPSERLQAMAEVPGLSAHPYLSG 60
QY 61 VERACFMSHVAVMKQALDBGLPYITVFEDVVLGSEBEKFLADAWLQRPDDTAPIYR 120
DB 61 VERACFMSHVAVMKQALDBGLPYITVFEDVVLGSEBEKFLADAWLQRPDDTAPIYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPPLSEBHWGTAGYIISRKAMRFLDRFALPPEGGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLSEBHWGTAGYIISRKAMRFLDRFALPPEGGLHP 180
QY 181 VDIAMFSDPFDRGMPVQCNLPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDIAMFSDPFDRGMPVQCNLPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
QY 241 PANTFGRRLIRALTXTISRERKRRORRQ----FTYFPQ 275
DB 241 PANTFGRRLIRALTXTISRERKRRORRQ----FTYFPQ 275

RESULT 5
US-11-102-497-8
Sequence 8, Application US/11102497
GENERAL INFORMATION:
APPLICANT: Gotschlich, Emil C.
TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND
TITLE OF INVENTION: GENES ENCODING THEM
FILE REFERENCE: 040853-01-5029-03
CURRENT APPLICATION NUMBER: US/11/102,497
CURRENT FILING DATE: 2005-04-08
PRIOR APPLICATION NUMBER: US/10/654,528
PRIOR FILING DATE: 2003-09-02
PRIOR APPLICATION NUMBER: US 10/007,267
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: US 09/333,412
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: US 08/878,360
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: US 08/683,426
PRIOR FILING DATE: 1996-07-18
PRIOR APPLICATION NUMBER: US 08/312,387
PRIOR FILING DATE: 1994-09-24
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.2

SEQ ID NO 8
LENGTH: 279
TYPE: PR1
ORGANISM: Neisseria gonorrhoeae
US-11-102-497-8

Query Match 89.1%; Score 1288; DB 41; Length 279;
Best Local Similarity 90.0%; Pred. No. 4,2e-121;
Matches 251; Conservative 7; Mismatches 17; Indels 4; Gaps 1;

QY 1 MONHVLSLSAARRRAHIAIDTFRHGIPPOFPDAMPSERLQAMAEVPGLSAHPYLSG 60
DB 1 MONHVLSLSAARRRAHIAIDTFRHGIPPOFPDAMPSERLQAMAEVPGLSAHPYLSG 60
QY 61 VERACFMSHVAVMKQALDBGLPYITVFEDVVLGSEBEKFLADAWLQRPDDTAPIYR 120
DB 61 VERACFMSHVAVMKQALDBGLPYITVFEDVVLGSEBEKFLADAWLQRPDDTAPIYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPPLSEBHWGTAGYIISRKAMRFLDRFALPPEGGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLSEBHWGTAGYIISRKAMRFLDRFALPPEGGLHP 180
QY 181 VDIAMFSDPFDRGMPVQCNLPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDIAMFSDPFDRGMPVQCNLPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
QY 241 PANTFGRRLIRALTXTISRERKRRORRQ----FTYFPQ 275
DB 241 PANTFGRRLIRALTXTISRERKRRORRQ----FTYFPQ 275

RESULT 6
US-10-467-657-1310
Sequence 1310, Application US/10467657
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACT Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqWin99, version 1.04
SEQ ID NO 1310
LENGTH: 279
TYPE: PR1
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1310

Query Match 88.6%; Score 1281; DB 34; Length 279;
Best Local Similarity 90.0%; Pred. No. 2,2e-120;
Matches 251; Conservative 6; Mismatches 18; Indels 4; Gaps 1;

QY 1 MONHVLSLSAARRRAHIAIDTFRHGIPPOFPDAMPSERLQAMAEVPGLSAHPYLSG 60
DB 1 MONHVLSLSAARRRAHIAIDTFRHGIPPOFPDAMPSERLQAMAEVPGLSAHPYLSG 60
QY 61 VERACFMSHVAVMKQALDBGLPYITVFEDVVLGSEBEKFLADAWLQRPDDTAPIYR 120
DB 61 VERACFMSHVAVMKQALDBGLPYITVFEDVVLGSEBEKFLADAWLQRPDDTAPIYR 120
QY 121 LETMFHVLTPSPGVADYCGRAFPPLSEBHWGTAGYIISRKAMRFLDRFALPPEGGLHP 180
DB 121 LETMFHVLTPSPGVADYCGRAFPPLSEBHWGTAGYIISRKAMRFLDRFALPPEGGLHP 180
QY 181 VDIAMFSDPFDRGMPVQCNLPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240
DB 181 VDIAMFSDPFDRGMPVQCNLPALCAQELHYAKFHDONSALGSLIEHDLRLNRKQQRDS 240

Db 61 VEKACPMASHAVLMKQALDEGLPYITVFEEDVLLGKDAEKFALBDTWLEERFDKDSAFIVR 120
Qy 121 LETMFMHVLTSBPSGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
Db 121 LETMFAKIVRPDKVLTNYENRSFPLSEHMGTAGYIISRREARFFLDRAVLPERRIKR 180
Qy 181 VDLMMFSDPDRGMPVCOINPALCAQELHYAKFHDONSALGSLIEHDLNRKQOORDS 240
Db 181 VDLMMFTYFDKGMPIYVAVSALCTQELHYAKFSLQNSMLGSDLEKD---REQRRRR 236
Qy 241 PANTFKHRLIRALTGISREKRRORRQ 269
Db 237 RSLKVMFDLKRALGKFGREKRRORRQ 265

RESULT 13
US-10-472-260-176 *
/ Sequence 176, Application US/10472260
/ GENERAL INFORMATION:
/ APPLICANT: MICROBIOLOGICAL RESEARCH AUTHORITY
/ TITLE OF INVENTION: IMPERIAL COLLEGE INNOVATIONS LIMITED
/ FILE REFERENCE: GMS/DJC/23480
/ CURRENT APPLICATION NUMBER: US/10/472,260
/ NUMBER OF SEQ ID NOS: 199
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 176
/ LENGTH: 280
/ TYPE: PRF
/ ORGANISM: Neisseria meningitidis (group B)
US-10-472-260-176

Query Match 64.5%; Score 933; DB 34; Length 280;
Best Local Similarity 69.5%; Pred. No. 4.1e-85;
Matches 187; Conservative 24; Mismatches 54; Indels 4; Gaps 1;

Qy 1 MONHVISLASAARRAHIAIDTFGRHGIPPOFDALMPSERLBOAMALVPGLSAHFYLSG 60
Db 1 MONHVISLASAARRAHIAIDTFGRHGIPPOFDALMPSERLBOAMALVPGLSAHFYLSG 60
Qy 61 VEKACPMASHAVLMKQALDEGLPYITVFEEDVLLGKDAEKFALBDTWLEERFDKDSAFIVR 120
Db 61 VEKACPMASHAVLMKQALDEGLPYITVFEEDVLLGKDAEKFALBDTWLEERFDKDSAFIVR 120
Qy 121 LETMFMHVLTSBPSGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 180
Db 121 LETMFAKIVRPDKVLTNYENRSFPLSEHMGTAGYIISRREARFFLDRAVLPERRIKR 180
Qy 181 VDLMMFSDPDRGMPVCOINPALCAQELHYAKFHDONSALGSLIEHDLNRKQOORDS 240
Db 181 VDLMMFTYFDKGMPIYVAVSALCTQELHYAKFSLQNSMLGSDLEKD---REQRRRR 236
Qy 241 PANTFKHRLIRALTGISREKRRORRQ 269
Db 237 RSLKVMFDLKRALGKFGREKRRORRQ 265

RESULT 14
US-60-691-214-3189
/ Sequence 3189, Application US/60691214
/ GENERAL INFORMATION:
/ APPLICANT: Munson, et al.
/ TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF H. INFLUENZAE
/ FILE REFERENCE: 28335/38815
/ CURRENT APPLICATION NUMBER: US/60/691,214
/ NUMBER OF SEQ ID NOS: 4387
/ SOFTWARE: Patentin version 3.3
/ SEQ ID NO 3189
/ LENGTH: 312
/ TYPE: PRF
/ ORGANISM: H. influenzae strain 86-028NP

US-60-691-214-3189

Query Match 30.2%; Score 436; DB 50; Length 312;

Best Local Similarity 35.0%; Pred. No. 1.1e-34;

Matches 100; Conservative 58; Mismatches 110; Indels 18; Gaps 5;

Qy 1 MONHVISLASAARRAHIAIDTFGRHGIPPOFDAL-----MPSERLEQAMAE---LVP 50
Db 1 ENYIYSNMENRTERKHITTKQFESKLSFSPFNATYQSINQSINQSINQSINQSISILH 85
Qy 51 GLSAHPIYSGVEKACPMASHAVLMKQALDEGLPYITVFEEDVLLGKDAEKFALBDTWLE 110
Db 86 NIEERILLTKGEGGLISHFILMNKCVNENLEYITFEEDVLLGNAEVLFLQDEMLTKR 145
Qy 111 PDDPTAFIVRLTETMFMHVLTSBPSGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLD 169
Db 146 PDEFNIFILRIETFLQPVKLEKQTKIPPFYSNFDILKSTHMGTAGYIISQAAKVIEX 205
Qy 170 PAALPPEGGLHPVDLMMFSDPDRGMPVCOINPALCAQELHYAKFHDONSALGSLIEH 229
Db 206 LKNISDEIVAVDELIFNKLVDVNVYIVYQINPALCTQEL--QANQSKSVLTSGLEK 262
Qy 230 ILNRKQOORDSPANTFKHRLIRALTGISREKRRORRQFIVPQ 275
Db 263 EKRPKIRKK---KTLKQRLTRIKENITIRALNRKKMKKQORIKEMQ 304

RESULT 15

US-10-603-108-2407

/ Sequence 2407, Application US/10603108

/ GENERAL INFORMATION:

/ APPLICANT: Gary L. Breton

/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARR

/ FILE REFERENCE: PAT93-14

/ CURRENT APPLICATION NUMBER: US/10/603,108

/ CURRENT FILING DATE: 2003-06-24

/ PRIOR FILING DATE: 2000-04-04

/ PRIOR APPLICATION NUMBER: US 60/125,416

/ NUMBER OF SEQ ID NOS: 3840

/ SEQ ID NO 2407

/ LENGTH: 255

/ TYPE: PRF

/ ORGANISM: M. catarrhalis

US-10-603-108-2407

Query Match 25.2%; Score 364; DB 36; Length 255;

Best Local Similarity 35.2%; Pred. No. 1.7e-27;

Matches 90; Conservative 49; Mismatches 103; Indels 14; Gaps 7;

Qy 1 MONHVISLASAARRAHIAIDTFGRHGIPPOFDALMPSERLBOAMALVPGLSAHFYLSG 60
Db 3 IONFVIVSVTATKRRRHIMCEFGKQIAEFEDAVPTDISKYAQLSIPINORLTDG 62
Qy 61 VEKACPMASHAVLMKQALDEGLPYITVFEEDVLLGKDAEKFALBDTWLEERFDKDSAFIVR 120
Db 63 -EKACFLSHVALMQQIDENLDYMAIFEDDVILGNSQOFLQELTWTQO---NAVVI 117
Qy 120 RLETFMFMHVLTSBPSGVADYCGRAFPILSEHMGTAGYIISRKAMRFFLDRAALPPEGGLHP 179
Db 118 KLET-WEKIHITKKAATVLTNRHQLCPDKTFPHGTAGYVISOQAKIILDYLTLDLAEFF 176
Qy 180 PVDLMMFSDPDRGMPVCOINPALCAQELHYAKFHDONSALGSLIEHDLNRKQOORD 239
Db 177 PIDHVFIDALISK--MSVLQVNPAAVITQ---AHVSEBDTFSIETGRKQVNVQHRRR 230
Qy 240 SPANTFKHRLIRALT 255
Db 231 TLAD-VGKKYRSIGK 245

Mon Apr 10 08:45:28 2006

us-09-211-691-2.rapm

Page 7

Search completed: April 7, 2006, 15:41:07
Job time : 575 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 15:31:41 ; Search time 60 Seconds
(without alignments)
630.723 Million cell updates/sec

Title: US-09-211-691-2

Perfect score: 1446

Sequence: 1 MGNHYSLASAEARRAHAD.....ISREKRRQRREQIVPRQ 275

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 584368 seqs, 137612332 residues

Total number of hits satisfying chosen parameters: 584368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending Patents AA New:*
1: /SIDS5/ptodata/1/paa/PC9_NEW_COMB.pep:*
2: /SIDS5/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /SIDS5/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /SIDS5/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /SIDS5/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /SIDS5/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /SIDS5/ptodata/1/paa/US11_NEW_COMB.pep:*
8: /SIDS5/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130.5	9.0	622	8	US-60-742-219-2224 Sequence 2224, Ap
2	113.5	7.8	527	6	US-10-513-957-29 Sequence 29, Appl
3	110	7.6	739	6	US-10-461-673-16345 Sequence 16345, A
4	90	6.2	738	6	US-11-311-940-9848 Sequence 9848, Ap
5	89	6.2	475	7	US-11-360-355-121052 Sequence 121052,
6	88.5	6.1	651	8	US-60-752-355-26732 Sequence 26732, A
7	87	6.0	771	7	US-11-360-355-133331 Sequence 133331,
8	85	5.9	1189	8	US-60-752-355-52377 Sequence 52377, A
9	84.5	5.8	386	1	PCT-US06-00964-4615 Sequence 4615, Ap
10	84.5	5.8	386	6	US-11-330-403-4815 Sequence 4815, Ap
11	84.5	5.8	951	8	US-60-752-355-16259 Sequence 16259, A
12	84	5.8	611	7	US-11-360-355-146612 Sequence 146612,
13	83.5	5.8	115	6	US-10-570-220-82 Sequence 82, Appl
14	83.5	5.8	479	8	US-60-752-355-26691 Sequence 26691, A
15	83.5	5.8	531	6	US-11-293-697-3858 Sequence 3858, Ap
16	83.5	5.8	531	8	US-60-751-420-2724 Sequence 2724, Ap
17	83.5	5.8	653	6	US-11-045-004-468 Sequence 468, App
18	83.5	5.8	803	6	US-11-045-004-2081 Sequence 2081, Ap
19	83.5	5.8	1217	6	US-10-461-673-9202 Sequence 9202, Ap
20	83.5	5.8	1336	6	US-60-742-219-1084 Sequence 1084, Ap
21	82.5	5.7	275	6	US-11-311-940-8938 Sequence 8938, Ap
22	82.5	5.7	873	8	US-60-752-355-39053 Sequence 39053, A
23	82.5	5.7	924	8	US-60-752-355-3969 Sequence 3969, Ap
24	82.5	5.7	966	8	US-60-752-355-1007 Sequence 1007, Ap
25	82.5	5.7	966	8	US-60-752-355-24368 Sequence 24368, A

26	82.5	5.7	966	8	US-60-752-355-38250 Sequence 38250, A
27	82	5.7	120	6	US-11-290-687-61 Sequence 61, Appl
28	82	5.7	955	8	US-60-752-355-23148 Sequence 23148, A
29	82	5.7	963	8	US-60-752-355-23441 Sequence 23441, A
30	82	5.7	1014	8	US-60-752-355-5751 Sequence 5751, Ap
31	81.5	5.6	316	6	US-10-461-673-10850 Sequence 10850, A
32	81.5	5.6	326	6	US-11-311-940-1350 Sequence 1350, Ap
33	81.5	5.6	326	8	US-60-752-355-41674 Sequence 41674, A
34	81.5	5.6	492	6	US-11-311-940-2862 Sequence 2862, Ap
35	81.5	5.6	492	8	US-60-752-355-17769 Sequence 17769, A
36	80.5	5.6	275	6	US-11-311-940-4346 Sequence 4346, Ap
37	80.5	5.6	297	6	US-11-311-940-6589 Sequence 6589, Ap
38	80.5	5.6	297	6	US-11-311-940-551 Sequence 551, App
39	80.5	5.6	312	6	US-11-311-940-5403 Sequence 5403, Ap
40	80.5	5.6	473	7	US-60-752-355-33231 Sequence 33231, A
41	80	5.5	530	7	US-11-360-355-124908 Sequence 124908, A
42	80	5.5	676	8	US-60-752-355-24908 Sequence 24908, A
43	80	5.5	1583	8	US-60-752-355-17545 Sequence 17545, A
44	79.5	5.5	492	1	PCT-US06-00964-5647 Sequence 5647, Ap
45	79.5	5.5	492	6	US-11-311-940-4124 Sequence 4124, Ap

ALIGNMENTS

RESULT 1
US-60-742-219-2224
; Sequence 2224, Application US/60742219
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John Verner
; APPLICANT: Bradley, Walter Edward
; APPLICANT: Paquin, Bruno
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Little, Randall David
; APPLICANT: Keith, Tim
; APPLICANT: Cousineau, Johane
; APPLICANT: Berdewegh, Paul Van
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
; FILE REFERENCE: 306522-2000
; CURRENT APPLICATION NUMBER: US/60/742,219
; NUMBER OF SEQ ID NOS: 7303
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2224
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-742-219-2224

Query Match 9.0%; Score 130.5; DB 8; Length 622;

Best Local Similarity 25.1%; Pred. No. 0.00696; Matches 58; Conservative 37; Mismatches 83; Indels 53; Gaps 12;

QY 5 VISLASAARRAHADTGRGHGIPQFPDA---MPSRLQMAALVPGISAPRY--- 57
DB 345 MINLRQRDRRRRLRALQAQIBCRALVAVDVKRAMTSQVVALGIQMLPGV-RDPYHGR 403
QY 58 -LSGVKACFSMNAVLMQALDEGLPIYTVPEDDV-----LIGSEERFLAED 104
DB 404 PLTKBELQCFISHYIMKRYVDKGLQKSLVFEDDDILFEIFKRRMLNMRDVE----- 456
QY 105 AWLQRRPDPVAFIVRLTMEHVLVTSQVADYCGRAP---LLSESH-WGTAGYIIS 159
DB 457 ---RGGLMDLIVYRKRMQVZH---PE-----KAVPRVNLVEADYSTWTLAYIIS 502
QY 160 RKAAMPFLDRFAALPPEGALPVDLMFSDPFDRECMPCQALNPAACQELH 210
DB 503 LQGAARKL---AABPLSKMLPVDKFL-PVMEFKH---PVSEYKAHPSRLNLIH 547

```
RESULT 2
US-10-513-957-29
; Sequence 29, Application US/10513957
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: KHARE, Reena
; APPLICANT: ELIJOT, Vicki S.
; APPLICANT: MARQUIS, Joseph P.
; APPLICANT: RAMKUMAR, Jayalakshi
; APPLICANT: CHAMPA, Narinder K.
; APPLICANT: MASON, Patricia M.
; APPLICANT: HAPALIA, April J.A.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: JIN, Pei
; APPLICANT: BECHA, Shanya D.
; APPLICANT: KABLE, Amy E.
; APPLICANT: TRAN, Oyen K.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BUREORD, Neil
; APPLICANT: GRAULI, Richard C.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: SPRAGUE, William W.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: CELL ADHESION AND EXTRACELLULAR MATRIX PROTEINS
; FILE REFERENCE: PF-1415 PCT
; CURRENT APPLICATION NUMBER: US/10/513,957
; PRIOR FILING DATE: 2004-11-10
; PRIOR APPLICATION NUMBER: US 60/379,840
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: US 60/381,291
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/383,183
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/394,146
; PRIOR FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PERL Program
; SEQ ID NO 29
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7512909CD1
US-10-513-957-29

Query Match      7.8%; Score 113.5; DB 6; Length 527;
Best Local Similarity 32.6%; Pred. No. 0.03;
Matches 31; Conservative 16; Mismatches 39; Indels 9; Gaps 3;

QY      5 VISLSAABRAHIAITPGRHGIPQFPFA---LMPSERLBQAMAEIVPGLSAHFY--- 57
DB      324 VLSLARPRRERRLASLMEISGRVDAVDGMLNSAIRNLGVLLPGIQ-DPYSGR 382
QY      58 -LSGVEKACFMSHAVLWKOALDEGLPYITVFEDDV 91
DB      383 TLTKEVGCFLSHYSIWEEVVARGLARVLVFEDDV 417

RESULT 3
US-10-461-673-16345
; Sequence 16345, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Aunudi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.

; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Ma, Yundong
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Ghosh, Malabika
; APPLICANT: Weng, Gezhi
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 10/243,552
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
; SOFTWARE: perl_genes Version 6.0
; SEQ ID NO 16345
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-461-673-16345

Query Match      7.6%; Score 110; DB 6; Length 739;
Best Local Similarity 24.0%; Pred. No. 0.099;
Matches 46; Conservative 34; Mismatches 82; Indels 30; Gaps 8;

QY      38 SERLSQAMAEIVPG---LSAHPYLSGVEKACFMSHAVLWKOALDEGLPYITVFEDDVLL 93
DB      496 TSQLRANINEMLPGYRDPSSRLPRG--BIGGFLSHYSVWKEVIDRELEKTLVIDDVAF 554
QY      94 GEGEEKFLAE--DANLQERFPDPTAFIVRLTFMFMHVLTPSGVADYCGRAFLLESSEH- 150
DB      555 EHQFFKKMKMLKMDNIDQALDWELIYIGRKMQVKEPEKAVENVAN-----LVEADVS 607
QY      151 WGTAGYIISRKAMRFFLDRFALPPEGHPVDLMM-----FSPDFDRGMPVQC 199
DB      608 YWTLGIYVLSLGAQQLV--GANPFGKMLPVDEFIPVMTNKPVAIEYKEIYESSRLDKAFS 664
QY      200 INPALCAOELAY 211
DB      665 AEP-LLIYPTHY 675

RESULT 4
US-11-311-940-9848
; Sequence 9848, Application US/11311940
; GENERAL INFORMATION:
; APPLICANT: Adad, Mark et al
```

```

: TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
:
: PILE REFERENCE: 38-21(53720)
: CURRENT APPLICATION NUMBER: US/11/311,940
: CURRENT FILING DATE: 2005-12-19
: NUMBER OF SEQ ID NOS: 10032
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 9848
:
: LENGTH: 738
: TYPE: prt
: ORGANISM: Chlamydia trachomatis D/UW-3/CX
US-11-311-940-9848

```

Query Match	6.2%; Score 90; DB 6; Length 738;
Best local similarity	33.4%; Score 140; DB 7.4;

```

Qy 67 MSHAVLMKQALDEGL-----PYITVF-----EDDVLLGSGEEKFLAEDAMLOERF--- 111
Db 184 K-----MEWVTSGGVLLIKSDPYGKGFKEPPWSVSVIDDSYE--WTDSEMLREERTKTE 236
Qy 112 PDPTAFIYRL-----ETMFEMVLTSPSGADVCGR-----APFLPLB--SEHNG--T 153
Db 237 GPMNIYEHAVGSRMRQEGOPRIYKELADQALCYCKQMHYTHVALLVYTEHPLNWSNGYOT 296
Qy 154 AGYII-----SRKAMRFPLDRFAA-----LPPEGIHPVDLMMFSDPFDRBCGVPVC 198
Db 297 TGYUAPTSGRYSGFEDLOYFIDTMHQHIGIVILDMWB--GHFPIDSPAMSGF--DGTPLY 351
Qy 199 Q-----LNPALCAQELHYAKPHDNNSALSLI 225
Db 352 EYTRNPSPLHPHMTYTFDYAKREYCNFLLGSVL 385

```

RESULT 5
 US-11-360-355-121052
 Sequence 121052, Application US/11360355
 GENERAL INFORMATION:
 APPLICANT: Boukharov, Andrey
 APPLICANT: Du, Zijiang
 APPLICANT: Guo, Ziliang
 APPLICANT: Kovalic, David
 APPLICANT: Lu, Maolong
 APPLICANT: McCarter, James
 APPLICANT: Miller, Nancy
 APPLICANT: Williams, Deryck
 APPLICANT: Vaudin, Mark
 APPLICANT: Wu, Wei
 TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
 TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
 FILE REFERENCE: 38-21(53885)
 CURRENT APPLICATION NUMBER: US/11/360,355
 CURRENT FILING DATE: 2006-02-24
 NUMBER OF SEQ ID NOS: 171306
 SEQ ID NO 121052
 LENGTH: 475
 TYPE: PRT
 ORGANISM: Heterodera glycyines
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (92)_(192)
 OTHER INFORMATION: Xaa can be any naturally occurring amino acid
 FEATURE:
 OTHER INFORMATION: Coding regions on vCDNA: vCDNA=SeqID 47475; Strand=-; Position=1
 OTHER INFORMATION: -63,173-550,843-1061,1104-1187,1230-1552,1593-1686,1782-1819
 FEATURE:
 OTHER INFORMATION: Homolog annotation: Hit ID=CAE71465.1; Match Level="QueryCoverage=89%, HitCoverage=76%, E-value=8e-61, Identity=33%", Hit description="Hypothetical protein CBG18383 [Caenorhabditis briggsae]"
 FEATURE:

OTHER INFORMATION: Essential gene: C elegans homolog=D2045.9; phenotype=Gro, Rnp, Strp
US-11-360-355-121052

Query Match	6.28; Score 89; DB 7; Length 475;
-------------	-----------------------------------

```
QY 2 QNHNSLSASAERAAHTADTFGRGIGPQGFADMLPSERLQAMAEI--VGLSAHPY-- 57
      ||::||::: : : : ||::||::: : : : ||::||:::
Db 303 QVTILINLRPRPQKKEKKRBEVLHVLGIDPTTFHEAVDGNLTTHETSAHFLPGGF-DPIYK 361
      ||::||::: : : : ||::||::: : : : ||::||:::
QY 58 --LSGVEKACFMSHAYLWKQALDEGLPYITVPEDDVLGEGEKFLAEDAWLQERFDDPT 115
      ||::||::: : : : ||::||::: : : : ||::||:::
Db 362 RPKMRGGITGCFLSHFKIWEELVRNGHQRAVVLLEDDV-----RPSRNG 403
      ||::||::: : : : ||::||::: : : : ||::||:::
QY 116 AFIIVLEETM 124
      ||::: ||
Db 404 TLIIQ-KTM 411
```

```

RESULT 6
US-60-752-355-26732
; Sequence 26732, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OR INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ. ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26732
; LENGTH: 651
; TYPE: PR1
; ORGANISM: Brucella melitensis biovar Abortus
US-60-752-355-26732

```

Query Match	Score	DB	Length
6.1%	88.5	8	651

```

QY      DDVLLIGSGEKEFTLAEADAMLOGRFPDPDAFIVRLETFMFWLTPSGVADYCGRAE---PL 145
      89  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      DDGILIGSGEFSGAEFRERIED-YPPD---ELEKM---RFAAASADVIGRAHEKOPT 347
      299 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      L-----ESEH---WGTAcyLLISRRKMRFFLDRAALP-----DEGLHPRVLLM 184
      146 |-----|-----|-----|-----|-----|-----|-----|
Db      LIYWGGEIVHFRAGSVSGF--TRNRLLEFPDRVLAMPILAENGFTGVALGALRGARPYVEI 405
      348 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      MFSDF-----FDREGNPVCQLNPALCAQELHYAKFPDQ--SALGSLIE 226
      185 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      MFGDFCEVAAQJANGISIKVRHMGDGFPPVIVRVRVSPHTGSGQHSQSPSALFGMPR 465
      406 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      HDRLINRKQQRDSRPATPFKH 247
      227 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      GMRVY-----SPTNAFDY 478
      466 |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 7
US-11-360-355-133331
Sequence 133331, Application US/11360355
GENERAL INFORMATION:
APPLICANT: Boukharov, Andrey
APPLICANT: Du, Zijing
APPLICANT: Guo, Jiliang
APPLICANT: Kovalic, David
APPLICANT: Lu, Maolong
APPLICANT: McCarter, James
APPLICANT: Miller, Nancy
APPLICANT: Williams, Deryck
APPLICANT: Vaudin, Mark
APPLICANT: Wu, Wei
TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS

```

```

; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 133331
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATRE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID_59754; Strand=+; Position=1
; OTHER INFORMATION: -135,193-600,837-1514,1692-2379,2422-2619,2663-2871
; FEATRE:
; OTHER INFORMATION: Homolog annotation: Hit ID=BAI49477.1; Match level="QueryCoverage
; OTHER INFORMATION: =49%, HitCoverage=27%, E-value=1e-10, Identity=22%"; Hit descript
; OTHER INFORMATION: =hypothetical protein 42.t00003 [Entamoeba histolytica HM-1.iMS3]
US-11-360-355-133331

Query Match
Best Local Similarity 26.1%; Score 87; DB 7; Length 771;
Matches 24; Conservative 17; Mismatches 37; Indels 14; Gaps 2;

QY 135 MCVQCLNPLCQOELHYAFHONSALGSLIEHDLNKKQQRSDSPANTFPHRLI---- 250
DB 579 LPTQAVEAYCADIRYANHHAEEDVQKAKHQQILERRORALDEBRO--RQKLSIWR 636
QY 251 -----RALTKISREKRRORRORQFIVPF 274
DB 637 DEDROWAIVEDVLAELKIRQKQKQIDIVEY 668

RESULT 8
US-60-752-355-52377
; Sequence 52377, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: Patent version 3.3
; SEQ ID NO 52377
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-60-752-355-52377

Query Match
Best Local Similarity 23.7%; Score 85; DB 8; Length 1189;
Matches 75; Conservative 35; Mismatches 116; Indels 90; Gaps 19;

QY 14 REAHADTGTGRI--PFOFDMAMPSERLQMAELVPLGSAHPYISGVEKACPMGHAVL 72
DB 176 KTAHASSP--EHGVSFMOAISLMP-----ALAEIGSGFPPEPDSMT--ITHAVM 223
QY 73 WKQALQ-----EGLYITVFEDVDVLGGEKEF-----LAEDAWLQERFPDPAFIVR 120
DB 224 GEAAFGISPADAEIMATLRTLTDDRM-----EKLCFAASLAKKAVAGEQLTLDIAY--- 275
QY 121 LETMFHVLTSPSGVA-----DYCGRAFPILBSEHWCTAGYIISRKAMRFLLR 169
DB 276 -DDIFLHCENAPAVVAHRRALDEEKI PYNSEALPWRPSBDFGRF--RAVSSSSAM--FFLGA 332
QY 170 PAALPEGLHPVDLMFSDPFD-----REGMPVCOLNPLACAGEL-- 209
DB 333 GNDYP--NLNPDY---DPPDTLIGIGARIFEMRIIRNITRRGIATTSL--PPVSATALAA 385
QY 210 --HYAKFHONSALGSLIEHDLNKKQQRSDSPAN-----TPKHRLIPALTKISR---E 259
DB 386 VARSGAHGRQPLDAVLIDPGCGAFENIGRLAI PAIKGI GIA TQGHQVSLVCGPRPLVE 445
QY 260 REKRRORRORQFIVPFQ 275
```

```

DB 446 QOKR-----FRIPRR 455

RESULT 9
PCT-US06-00964-4815
; Sequence 4815, Application PC/TUS0600964
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: PCT/US06/00964
; CURRENT FILING DATE: 2006-01-23
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 4815
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Rubrobacter xylophilus DSM 9941
PCT-US06-00964-4815

Query Match
Best Local Similarity 21.3%; Score 84.5; DB 1; Length 386;
Matches 47; Conservative 30; Mismatches 61; Indels 83; Gaps 10;

QY 13 ERRAHADTGTGRI--PFOFDMAMPSE-----RLQAMAEVPLGSAHPYISGV----- 61
DB 28 ERRERLKE-----QGVRLFDGAGDREPTDPRIRALLAEAVDEVSQYPSAGTRALREAF 83
QY 62 -----EKACPMGHAVLM-----KQALDEGLPYITVEDVL 92
DB 84 CGMMERRRGVELDPDAEVLPAAGSKKALFHAHLPELHHTHRRGVAYGTPGYVERGAL 143
QY 93 LGEGER--KFLAEDAWL--QERFPDPAFIVRLETMFHVLTSPSGVADYCGRAFPILLE 147
DB 144 PAGGALVRLVREBGFILPLEAVDPQ-----KRIIMLANTPHNTGA----- 186
QY 148 SEHWCTAGYIISRKAMRFLLRFAALPEGLHPVDLMFSD 188
DB 187 -----RAPRSYLEBAAFCRE--H--DILLFSD 210
```

```

RESULT 10
US-11-330-403-4815
; Sequence 4815, Application US/11330403
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 4815
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Rubrobacter xylophilus DSM 9941
US-11-330-403-4815

Query Match
Best Local Similarity 21.3%; Score 84.5; DB 6; Length 386;
Matches 47; Conservative 30; Mismatches 61; Indels 83; Gaps 10;

QY 13 ERRAHADTGTGRI--PFOFDMAMPSE-----RLQAMAEVPLGSAHPYISGV----- 61
DB 28 ERRERLKE-----QGVRLFDGAGDREPTDPRIRALLAEAVDEVSQYPSAGTRALREAF 83
QY 62 -----EKACPMGHAVLM-----KQALDEGLPYITVEDVL 92
DB 84 CGMMERRRGVELDPDAEVLPAAGSKKALFHAHLPELHHTHRRGVAYGTPGYVERGAL 143
QY 93 LGEGER--KFLAEDAWL--QERFPDPAFIVRLETMFHVLTSPSGVADYCGRAFPILLE 147
DB 144 PAGGALVRLVREBGFILPLEAVDPQ-----KRIIMLANTPHNTGA----- 186
```


Qy 148 SEHMGTAGYIISRKAMFPLDRPALPPEGHLPVDLMMESD 188
Db 187 -----RAPRSTYEEBAAPCR--H--DILTFPD 210

RESULT 11

US-60-752-355-16259
; Sequence 16259, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720) B
; CURRENT APPLICATION NUMBER: US/60/752,355
; NUMBER OF FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16259
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (190)..(190)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-60-752-355-16259

Query Match 5.8%; Score 84.5; DB 8; Length 951;
Best Local Similarity 22.6%; Pred. No. 34;
Matches 85; Conservative 43; Mismatches 103; Indels 145; Gaps 23;

Qy 3 NHVISLASAAE-----RRAHITAD-----TFGR--HGI---PFOFF 32
Db 62 SHMLNLANIAEIQMYRRKKAESRRGDPVDEASAPTESDIDETFORLYRGKSGSPREVF 121
Qy 33 DALMPSERLQAMALVPGLSAHPYLSGVEKACFMSHAVALMKQALDEGPPYITVFEDVYL 92
Db 122 DA-----LRSCQITDLY--LTAHPQS--VRSRLQGNASI--RSLCLN-----L 159
Qy 93 LSEG-----BEKFLABAMIQERDPPTAFIVRLTFMVLVSPSGVADYCGRAFLPLS 148
Db 160 CEGGSDNEBQET--DEALQREMEMCSSGRIR-----XILAAFK--TBEIRRTPTPTPD 209
Qy 149 EHWGTAGYIISRKAMFPLDRPALPPEGHLPVD-----LMMFGDFP- 190
Db 210 E-----MRAGSYFHTDTMNGVPKFLKRVDTLAKIIGIDERLPYAAPLIQFSSWNG 260
Qy 191 -DREGHP-----VQQLNPALCAQELHYAKFHDONSALGSLIEHRL-----LNKK 234
Db 261 GDRDGNRYTPRYTRDVCLL--ARMMAAHVYFSKADLMFELSMWRCNDELBARAMELHK 319
Qy 235 QQRR-----DSPANF-----KHRLI--RALKISREKRRQ----- 265
Db 320 SSRKAYKYIASLCTPAATFWKKISPEGRYRIILGVDRLKVTYTERAQLILSKGISIP 379
Qy 266 -----RRQFIVPFQ 275
Db 380 EDQTYNVQFLFPLE 395

RESULT 12

US-11-360-355-146612
; Sequence 146612, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, MaoJong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei

; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; NUMBER OF FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 146612
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on vCDNA: vCDNA=SeqID 73035; Strand=-; Position=1
; OTHER INFORMATION: -189,235-786,1093-1780,1822-2019,2063-2271
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=NP_958786.1; Match level="QueryCovera
; OTHER INFORMATION: =87%, HitCoverage=11%, E-value=3e-15, Identity=211%; Hit descrip
; OTHER INFORMATION: =plectin 1 isoform 11 [Homo sapiens] gb|AA85684.1| plectin 11 [

US-11-360-355-146612

Query Match 5.8%; Score 84; DB 7; Length 611;
Best Local Similarity 25.0%; Pred. No. 21;
Matches 23; Conservative 18; Mismatches 37; Indels 14; Gaps 2;

Qy 195 MPVQQLNPALCAQELHYAKFHDONSALGSLIEHRLNKKQORRDSPPANTFKHRLI----- 250
Db 419 LPTROAVEAYCADITRYANHAEDVQKATKEHQQLERERLBERO--RQKLSTWR 476
Qy 251 -----RALKISREKRRQRRQFIVPF 274
Db 477 DEDQWAVEDEPVLALEKIRQKEQIDVEY 508

RESULT 13

US-10-570-220-82
; Sequence 82, Application US/10570220
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; APPLICANT: represented by the Secretary of the Department of Health and
; APPLICANT: Human Services
; APPLICANT: Kashmiri, Syed V.S.
; APPLICANT: Schlom, Jeffrey
; APPLICANT: Padlan, Eduardo A.
; TITLE OF INVENTION: MINIMALLY IMMUNOGENIC VARIANTS OF SDR-GRAFTED HUMANIZED ANTIBODY
; FILE REFERENCE: CC49 AND THEIR USE
; CURRENT APPLICATION NUMBER: US/10/570,220
; CURRENT FILING DATE: 2006-02-28
; PRIOR APPLICATION NUMBER: 60/498,903
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: PCT/US04/028004
; PRIOR FILING DATE: 2004-08-27
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-570-220-82

Query Match 5.8%; Score 83.5; DB 6; Length 115;
Best Local Similarity 29.8%; Pred. No. 2.6;
Matches 37; Conservative 19; Mismatches 37; Indels 31; Gaps 9;

Qy 40 RLQQAMAEIV-PGLSAHPYLSGVEKAC-----FMASHVLTW-KQALDDEGLPYITVF---E 88
Db 3 QLQGSDAELVYKGS-----VKISCKASGTFPTDHAIHVKQKPEQGLEIYIGFSPON 55
Qy 89 DDVLIGSGEKEFLABDAMIQERFPDPTAFIVRLTFMVLVSPSGVADYCGRAFLPLS 148
Db 56 DDF--KYMERPKGR-ATLTADKSSSTAYVQ-----INSILTSBDSAVYFCTRS---LNM 102
Qy 149 EHWG 152

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 15:58:52 ; Search time 280.271 Seconds
(without alignments)
8315.447 Million cell updates/sec

Title: US-09-211-691-3

Perfect score: 41
Sequence: 1 cttgaggagtcatactgtaaa.....acaaatattgcgtttac 41

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBdb1.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_ov.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vl.*
14: gb_hcg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	70.7	1875	1 U04328	U04328 Neisseria m
2	29	70.7	4176	1 NMBSYN	X78068 N.meningitidis
3	29	70.7	4908	1 NMBSYN	M95053 Neisseria m
4	29	70.7	110000	1 AE002098_00	AE002098 Neisseria
5	29	70.7	349980	6 AX043922	AX043922 Sequence
6	28	68.3	687	1 NMU60146	U60146 Neisseria m
7	25.8	62.9	224302	14 AC103481	AC103481 Rattus no
8	25.4	62.0	401	6 AR271714	AR271714 Sequence
9	25.4	62.0	41	6 AR481803	AR481803 Sequence
10	25.4	62.0	41	6 AR527402	AR527402 Sequence
11	25.4	62.0	41	6 AR609681	AR609681 Sequence
12	25.4	62.0	41	6 AX934469	AX934469 Sequence
13	25.2	61.5	10029	1 AE015576	AE015576 Shewanella
14	25.2	61.5	301439	1 AE015576	AE015576 Clostridi
15	24.8	60.5	139325	8 AL589943	AL589943 Human DNA
16	24.8	60.5	165199	8 AC003973	AC003973 Homo sapi
17	24.8	60.5	252718	9 AC116106	AC116106 Mus muscu
18	24.6	60.0	59915	14 AC100984	AC100984 Mus muscu

C	19	24.6	60.0	110000	14 BX890561_1	Continuation (2 of
C	20	24.6	60.0	131209	8 AC126177	AC126177 Homo sapi
C	21	24.6	60.0	141017	9 AC126444	AC126444 Mus muscu
C	22	24.6	60.0	180932	9 AC131989	AC131989 Mus muscu
C	23	24.6	60.0	191606	9 AC126272	AC126272 Mus muscu
C	24	24.6	59.5	2966	6 CQ714591	CQ714591 Sequence
C	25	24.4	59.5	4207	8 AF061935	AF061935 Homo sapi
C	26	24.4	59.5	4356	8 BC022792	BC022792 Homo sapi
C	27	24.4	59.5	5984	8 AB018343	AB018343 Homo sapi
C	28	24.4	59.5	168499	8 AC092037	AC092037 Homo sapi
C	29	24.2	59.0	720	10 BV655397	BV655397 S216P6047
C	30	24	58.5	110000	1 CP000020_13	Continuation (14 o
C	31	24	58.5	174375	9 AC124422	AC124422 Mus muscu
C	32	24	58.5	241329	14 AC106206	AC106206 Rattus no
C	33	23.8	58.0	152095	9 AC154271	AC154271 Mus muscu
C	34	23.8	58.0	167612	14 AC119140	AC119140 Rattus no
C	35	23.8	58.0	180558	9 AC121840	AC121840 Mus muscu
C	36	23.8	58.0	204672	9 AC122831	AC122831 Mus muscu
C	37	23.8	58.0	212749	14 AC115435	AC115435 Rattus no
C	38	23.6	57.6	50083	8 AL359708	AL359708 Human DNA
C	39	23.6	57.6	110000	1 CR522870_25	Continuation (26 o
C	40	23.6	57.6	148965	14 AC136069	AC136069 Rattus no
C	41	23.6	57.6	180745	9 AC123033	AC123033 Mus muscu
C	42	23.6	57.6	196791	14 AC118298	AC118298 Rattus no
C	43	23.4	57.1	4842	9 BC065119	BC065119 Mus muscu
C	44	23.4	57.1	4993	9 AK122372	AK122372 Mus muscu
C	45	23.4	57.1	12037	1 AE013886	AE013886 Versinia

ALIGNMENTS

RESULT 1	U04328	1875 bp	DNA	linear	BC15-APR-1994
LOCUS	U04328				
DEFINITION	Neisseria meningitidis NMB CMP-N-acetylneuraminic acid synthetase				
ACCESSION	U04328				
VERSION	U04328.1	GI:460144			
KEYWORDS					
SOURCE	Neisseria meningitidis				
ORGANISM	Neisseria meningitidis				
REFERENCE	1 (bases 1 to 1875)				
AUTHORS	Swartley, D.S. and Stephens, D.S.				
TITLE	Identification of a genetic locus involved in the biosynthesis of N-acetyl-D-mannosamine, a precursor of the (alpha 2->8)-linked polysialic acid capsule of serogroup B Neisseria meningitidis				
JOURNAL	J. Bacteriol. 176 (5), 1530-1534 (1994)				
PubMed	8113198				
REFERENCE	2 (bases 1 to 1875)				
AUTHORS	Stephens, D.S.				
TITLE	Direct Submision				
JOURNAL	Submitted (13-DEC-1993) David S. Stephens, Emory University, Department of Medicine, 69 Butler Street, Atlanta, GA 30303, USA				
FEATURES	Location/Qualifiers				
source	1..1875				
	/organism="Neisseria meningitidis"				
	/mol_type="unassigned DNA"				
	/strain="NMB"				
	/isolate="N7"				
	/db_xref="taxon:487"				
	/note="serogroup B"				
gene	52..1185				
	/gene="synx"				
CDS	52..1185				
	/gene="synx"				
	/function="proposed to be involved in the biosynthesis of N-acetyl-D-mannosamine"				
	/codon_start=1				
	/evidence=experimental				
	/transl_table=11				

```
/product="synx"
/protein_id="AA17654.1"
/db_xref="GI:460145"
/translation="MKRLICTTGRADPGKLPKLLAYIENHPDLHLIVTGMHMT
YGRYKVTRENYQHTYLFNSQIOGEPGAVLNTTIFSRLSDEIEPDWMHIDRL
BALAGAAVAGLSRLVCHIEGELSGTVDSIRHSISKLSHLVANEQAVTRLVQMG
EKRRHIIHIGSPDLVMASTLPSELEVEYVGLPYENYGISMFHPTTEAHLMPOYA
AOPYKALISGONIIISIPNDTGESILOELKQSDKFIAPFPIRFEYFLVLLKHA
KFWGNSAGIREAPLYGVPSIDVGTOSNRHMGKSIHTDYEETKNIFDAIQACSLG
KPEADDTFNGDTRTSTERPAVINNPETMNVSAQKRFIDNL"
1189..1875
/gene="synb"
1189..1875
/gene="synb"
*/standard_name="CMP-NANA synthetase"
/codon_start=1
/translation="1"
/product="CMP-N-acetylneuraminic acid synthetase"
/protein_id="AA17655.1"
/db_xref="GI:460146"
/translation="MEKONIAVILARQNSKGLPKNLKRNNGISLHGTINNAISKC
PRLIVSTGGLIABEAKNGVEVLRPAELASPTASISGVIALETIGNSGVTL
LOPSPRTGAHIREAPFLPDEKIRGSVSAQPMHHPKLTLOINNGEYAPMRHLS
LEQPRQOLPOAFRPRGALYINDTASLIANNCFFIAPTKLYIMSHQSDSIDIDTELQO
AENILNHKS"
```

ORIGIN

```
Query Match      70.7%; Score 29; DB 1; Length 1875;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      13 TATGGAAAACAAATATGCGGTATAC 41
      |||||||||||||||||||
Db      1188 TATGGAAAACAAATATGCGGTATAC 1216
```

```
RESULT 2
NMBSYN      4176 bp      DNA      linear      BCT 18-APR-2005
LOCUS      N.meningitidis (group B) ctra, synb, synb, sync and
DEFINITION      N.meningitidis (group B) ctra, synb, synb, sync and
ACCESSION      X78068
VERSION      X78068.1 GI:530037
KEYWORDS      CMP-NeuNAC synthetase; ctra gene; sialyltransferase; syna gene;
SOURCE      Neisseria meningitidis serogroup B
ORGANISM      Neisseria meningitidis serogroup B
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
1
REFERENCE      1
AUTHORS      Ganguli,S., Zapata,G., Wallis,T., Reid,C., Boulnois,G., Vann,W.F.
and Roberts,I.S.
TITLE      Molecular cloning and analysis of genes for sialic acid synthetase
in Neisseria meningitidis group B and purification of the
meningococcal CMP-NeuNAC synthetase enzyme
JOURNAL      J. Bacteriol. 176 (15), 4583-4589 (1994)
PUBMED      8045888
REFERENCE      2 (bases 1 to 4176)
AUTHORS      Ganguli,S.
TITLE      Direct Submission
JOURNAL      Submitted (09-MAR-1994) S. Ganguli, University of Leicester, Dept
of Microbiology, PO Box 138, Medical Sciences Building, University
Road, Leicester LE1 9HN, UK M64389.
COMMENT      Related sequences: M57677 & M64389.
FEATURES
source
1..4176
/organism="Neisseria meningitidis serogroup B"
/mol_type="genomic DNA"
/isolate="NCTC8249"
/db_xref="taxon:491"
/complement(<1..657)
/gene="ctrA"
complement(<1..657)
```

gene
mRNA

gene
CDS

gene
CDS

gene
CDS

CDS

stem_loop

ORIGIN

Query Match

```
/gene="ctrA"
792..1925
/gene="syna"
792..1925
/gene="syna"
/codon_start=1
/translation="1"
/protein_id="CAA54982.1"
/db_xref="GI:530038"
/db_xref="GOA:O57141"
/db_xref="InterPro:IPR003331"
/db_xref="UniProt/TREMBL:O57141"
/translation="MKRLICTTGRADPGKLPKLLAYIENHPDLHLIVTGMHMT
YGRYKVTRENYQHTYLFNSQIOGEPGAVLNTTIFSRLSDEIEPDWMHIDRL
BALAGAAVAGLSRLVCHIEGELSGTVDSIRHSISKLSHLVANEQAVTRLVQMG
EKRRHIIHIGSPDLVMASTLPSELEVEYVGLPYENYGISMFHPTTEAHLMPOYA
AOPYKALISGONIIISIPNDTGESILOELKQSDKFIAPFPIRFEYFLVLLKHA
KFWGNSAGIREAPLYGVPSIDVGTOSNRHMGKSIHTDYEETKNIFDAIQACSLG
KPEADDTFNGDTRTSTERPAVINNPETMNVSAQKRFIDNL"
1929..2615
/gene="synb"
1929..2615
/gene="synb"
*/standard_name="CMP-NeuNAC synthetase"
/codon_start=1
/translation="1"
/product="CMP-NeuNAC synthetase"
/protein_id="CAA54983.1"
/db_xref="GI:530039"
/db_xref="GOA:POA027"
/db_xref="InterPro:IPR003329"
/db_xref="UniProt/Swiss-Prot:POA027"
/translation="MEKONIAVILARQNSKGLPKNLKRNNGISLHGTINNAISKC
PRLIVSTGGLIABEAKNGVEVLRPAELASPTASISGVIALETIGNSGVTL
LOPSPRTGAHIREAPFLPDEKIRGSVSAQPMHHPKLTLOINNGEYAPMRHLS
LEQPRQOLPOAFRPRGALYINDTASLIANNCFFIAPTKLYIMSHQSDSIDIDTELQO
AENILNHKS"
2616..3665
/gene="sync"
2616..3665
/gene="sync"
/codon_start=1
/translation="1"
/protein_id="CAA54984.1"
/db_xref="GI:530040"
/db_xref="GOA:Q7DDU0"
/db_xref="InterPro:IPR004144"
/db_xref="InterPro:IPR006014"
/db_xref="InterPro:IPR006190"
/db_xref="UniProt/TREMBL:Q7DDU0"
/translation="WONNEPRTGNRSVGNVHEPILICEIGINHGSLKTAPEWDA
INAGAVAHQHTIYEDMSDBAKOVIFGNADVSIYEIMERCALNREDEIKLKEYES
KMIFISTPFSKALRLQRMIDIPAKIGSGECNNYPLIKVASFGKPIITSGNAST
BSIKKSVISIRAGVVALHCTNITYPPYEBVLRGANNDSSEAFPDATIGSDHLD
NACAGAAVAGLSRLVCHIEGELSGTVDSIRHSISKLSHLVANEQAVTRLVQMG
EKRRHIIHIGSPDLVMASTLPSELEVEYVGLPYENYGISMFHPTTEAHLMPOYA
AOPYKALISGONIIISIPNDTGESILOELKQSDKFIAPFPIRFEYFLVLLKHA
KFWGNSAGIREAPLYGVPSIDVGTOSNRHMGKSIHTDYEETKNIFDAIQACSLG
KPEADDTFNGDTRTSTERPAVINNPETMNVSAQKRFIDNL"
3749..>4176
/codon_start=1
/translation="1"
/product="sialyltransferase"
/protein_id="CAA54985.1"
/db_xref="GI:530041"
/db_xref="GOA:O79AE7"
/db_xref="InterPro:IPR010866"
/db_xref="UniProt/TREMBL:O79AE7"
/translation="MKRLICTTGRADPGKLPKLLAYIENHPDLHLIVTGMHMT
YGRYKVTRENYQHTYLFNSQIOGEPGAVLNTTIFSRLSDEIEPDWMHIDRL
BALAGAAVAGLSRLVCHIEGELSGTVDSIRHSISKLSHLVANEQAVTRLVQMG
EKRRHIIHIGSPDLVMASTLPSELEVEYVGLPYENYGISMFHPTTEAHLMPOYA
AOPYKALISGONIIISIPNDTGESILOELKQSDKFIAPFPIRFEYFLVLLKHA
KFWGNSAGIREAPLYGVPSIDVGTOSNRHMGKSIHTDYEETKNIFDAIQACSLG
KPEADDTFNGDTRTSTERPAVINNPETMNVSAQKRFIDNL"
3773..3790
/notes="putative"
```

70.7%; Score 29; DB 1; Length 4176;

Best Local Similarity 100.0%; Pred. No. 20;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGAAAAAATAATTTGGGTTATAC 41
Db 1928 TATGAAAAAATAATTTGGGTTATAC 1956

RESULT 3

NM95053

LOCUS 4908 bp DNA linear BCT 04-Apr-2002

DEFINITION Neisseria meningitidis (siae), CMP-NeuNac synthetase (siae),

ACCESSION NM95053 M64289 M96562

VERSION M95053.1 GI:520732

SOURCE Neisseria meningitidis

ORGANISM Neisseria meningitidis

Neisseriaceae; Neisseria.

REFERENCE 1 (bases 2861 to 4641)

AUTHORS Prosch,M., Edwards,U., Bousset,K., Krause,B. and Weisgerber,C.

TITLE Evidence for a common molecular origin of the capsule gene loci in

JOURNAL Mol. Microbiol. 5 (5), 1251-1263 (1991)

PUBMED 1659649

REFERENCE 2 (bases 1210 to 1987)

AUTHORS Edwards,U. and Prosch,M.

TITLE Sequence and functional analysis of the cloned Neisseria

JOURNAL FEMS Microbiol. Lett. 75 (2-3), 161-166 (1992)

PUBMED 1398032

REFERENCE 3 (bases 1 to 4908)

AUTHORS Edwards,U., Muller,A., Hammerschmidt,S. and Prosch,M.

TITLE Molecular analysis of the biosynthesis pathway of the alpha 2,8

JOURNAL polygalactonic acid capsule by Neisseria meningitidis serogroup B

PUBMED Unpublished

REFERENCE 4 (bases 1210 to 1987)

AUTHORS Prosch,M.

TITLE Direct Submission

JOURNAL Submitted (28-OCT-1992) Matthias Prosch, Medizinische Hochschule

Hannover, Institut für Medizinische Mikrobiologie, Konstanty

Gutschow Strasse 8, Hannover, Germany, 30623

COMMENT On or before Apr 4, 2002 this sequence version replaced gi:150354,

gi:497288.

FEATURES Location/Qualifiers

1..4908

/organism="Neisseria meningitidis"

/mol_type="genomic DNA"

/isolate="B1940"

/specific_host="Homo sapiens"

/db_xref="taxon:487"

51..56

/standard_name="-10 signal"

174..1307

/gene="siae"

174..1307

/gene="siae"

/codon_start=1

/transl_table=11

/product="Siae"

/protein_id="AAA20475.1"

/db_xref="GI:520732"

/translation="MKRIICITGTRADFGKLPKLAIVENHPDLIELIVTGMMNKT

YGRTRYEVRENQHTYLVFENIQGSPGAVLQNTITPISRLSDRIEPPDMVTHGRL

BALAGAVALISRLVCHIEGELSGVDSISHSISKSHILVANEOAVTLLVQMG

EKKHIIIGSPDLVDWASSTPLSLEVEKYGLPYENYGISMFHVTTBAHLMPPQA

gene

CDS

1311..1997

/gene="siae"

1311..1997

/gene="siae"

/citation=[2]

/codon_start=1

/transl_table=11

/product="CMP-NeuNac synthetase"

/protein_id="AAA20476.1"

/db_xref="GI:520734"

/translation="MEKONIAVILARONSGPLIKYLRRMNGISLGHITNAISKC

FRRITVSTGGLIAEBAKFGVYVLRPAELASDPTASSISGVTHALETIGNSGTYTL

LQPTSPRTGAHREAFSLPDEKIGSVSACWEPHPKTLQINNGEAPRHLSD

LQOPRQOLPQAFRPNQAIYINDYASLIANNCFPIATPKYIMSHQSDIDIDELDQ

AENILNHKS"

1998..3047

/gene="siae"

1998..3047

/gene="siae"

/citation=[1]

/codon_start=1

/transl_table=11

/product="siae"

/protein_id="AAA20477.1"

/db_xref="GI:520735"

/translation="MQNNEPFIKGRSVGVNHPPLICIGINHGSLKTAPEMVA

YNAGAEVGHQTHIVDEMSDEAKQVTPNAVYSIYMERCLNEDSILKAEYES

KCMIFSTPSPSAAALRIORDIIPAKISGEONNPLIKVASPEKPIIISGNSMI

BSIKSVETIPRRAGVYVLRPAELASDPTASSISGVTHALETIGNSGTYTL

NYACGVALGSSILERHFTDMRPGDIVCSMPDTKELQGHAKLARGKKD

TIIAGKPTKDPAPASVADKDIKGGELSGNLWKRNGDPSVNEYETLFGKVA

CNIRKGAQIKTKDIE"

3131..4618

/gene="siae"

3131..4618

/gene="siae"

/citation=[3]

/codon_start=1

/transl_table=11

/protein_id="AAA20478.1"

/db_xref="GI:520736"

/translation="MLKIKKALPQPKKFFQDSMWLTSPFYLTTPRNNLFEVINSQ

LQVOSLIRIKLITNLIVLYTSKRLKMPKLVHOSANRNFESILFELPRSPNIT

PKKLIVYTSYKKNINIOPALVMSFTGHSYLSIAKKNIITHLIDEGTYAP

LLESFSPYTKLERYLIGNLNIKGYIDHFDILHVPFPAKKIIPAKKYNRPAHAG

GISINNINLAKKYOISKNQYIPVSOYRPIISDLYKRSIVITINSISLOIGKFIK

LHPKKNRYVMSLPLNWEINPRIVINPEPLIFPLIYLTNRKQITGLASSLIYT

PLSPSTGLSTIGELTINLQKTSVNTENTMIOEHLEITIKRNFILINDLNQVSNP

LPTBETFEFTLLKSAEPAYKSNYQALFYQGLASKNITTLGHKALWYNALYVNO

IYMEYSIDIFYIDNISVDFHSKDLTKWIKIKHYYSADNRIGDR"

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 19;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGAAAAAATAATTTGGGTTATAC 41
Db 1310 TATGAAAAAATAATTTGGGTTATAC 1338

RESULT 4

AE002098 00/c

Sequence split into 23 fragments LOCUS AE002098 Accession AE002098

Fragment Name Begin End

AE002098_00 1 110000

AE002098_01 100001 210000

AE002098_02 200001 310000

AE002098_03 300001 410000

	AE002098_04	400001	510000	
	AE002098_05	500001	610000	
	AE002098_06	600001	710000	
	AE002098_07	700001	810000	
	AE002098_08	800001	910000	
	AE002098_09	900001	1010000	
	AE002098_10	1000001	1110000	
	AE002098_11	1100001	1210000	
	AE002098_12	1200001	1310000	
	AE002098_13	1300001	1410000	
	AE002098_14	1400001	1510000	
	AE002098_15	1500001	1610000	
	AE002098_16	1600001	1710000	
	AE002098_17	1700001	1810000	
	AE002098_18	1800001	1910000	
	AE002098_19	1900001	2010000	
	AE002098_20	2000001	2110000	
	AE002098_21	2100001	2210000	
	AE002098_22	2200001	2272360	
LOCUS	AE002098	2272360 bp	DNA	circular BCT 26-MAY-2005
DEFINITION	Neisseria meningitidis MC58, complete genome.			
ACCESSION	AE002098 AE002359-AE002564			
VERSION	AE002098.2 GI:66731897			
KEYWORDS				
SOURCE				
ORGANISM	Neisseria meningitidis MC58			
	Neisseria meningitidis MC58			
	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;			
	Neisseriaceae; Neisseria.			
REFERENCE	1 (bases 1 to 2272360)			
AUTHORS	Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Cickco,A., Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T.M., Parksey,D.S., Blair,E., Clifton,H., Clifton,H., Clark,E.B., Cotton,M.D., Utecherack,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scariato,V., Maignani,V., Pizsa,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappuoli,R. and Venter,J.C.			
TITLE	Complete genome sequence of Neisseria meningitidis serogroup B strain MC58			
JOURNAL	Science 287 (5459), 1809-1815 (2000)			
PUBLISHED	10710307			
REFERENCE	2 (bases 1 to 2272360)			
AUTHORS	Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Cickco,A., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T.M., Parksey,D.S., Blair,E., Clifton,H., Clark,E.B., Cotton,M.D., Utecherack,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scariato,V., Maignani,V., DeBoy,R.T., Pizsa,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappuoli,R. and Venter,J.C.			
TITLE	Direct Submission			
JOURNAL	Submitted (17-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			
REFERENCE	3 (bases 1 to 2272360)			
AUTHORS	Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Cickco,A., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T.M., Parksey,D.S., Blair,E., Clifton,H., Clark,E.B., Cotton,M.D., Utecherack,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scariato,V., Maignani,V., DeBoy,R.T., Pizsa,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappuoli,R. and Venter,J.C.			
TITLE	Direct Submission			
JOURNAL	Submitted (18-MAY-2005) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			
REFERENCE	Sequence update by submitter			
REMARK	On or before May 26, 2005 this sequence version replaced			
COMMENT	gi:7225225 gi:7225241 gi:7225245 gi:7225252 gi:7225264 gi:7225269 gi:7225271 gi:7413418 gi:7413419 gi:7225303 gi:7225327 gi:7225337 gi:7413420 gi:7413421 gi:7225314			

FEATURES

source

1. .227360
/organism="Neisseria meningitidis MCS8"
/mol_type="genomic DNA"
/strain="MCS8"
/db_xref="taxon:122586"
/note="serogroup: B"

gene

/locus_tag="NMB0001"
complement (7. .498)
/locus_tag="NMB0001"
/note="identified by similarity to EGAD.160712; match to protein family HMM PF00583"

CDS

/codon_start=1
/transl_table=11
/product="putative acetyltransferase"
/protein_id="AAF0480.1"
/db_xref="GI:7225226"
/translation="MNSTLYPDYNTVITRLKAGHIGRLVQALPRMGHFEPMSVDKIH
AYGRCKLDEPLAPRAVDDSGILGASNAVKXHDHESFRITYGLADVVLPEYRKK
GIGRLVAHCIGAARSIGIFLTYLPDVDFESREWVVGGHFNHGAVTWRLDVG
DKV"

gene

complement (502. .897)
/locus_tag="NMB0002"
complement (502. .897)
/locus_tag="NMB0002"
/note="identified by glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF0481.1"
/db_xref="GI:7225227"
/translation="MPSDVGIRLOTAFAFKMLRKKKIIFYPLMVVFSTSVAGDAEDNL
SISGVALLQKNNIDRGKITGNOSDIEDRRRLQTRQADIARLEAIPANMKAKAQ

gene
complement (918..2312)
/gene="glex"
/locus_tag="NMB0003"
complement (918..2312)
/gene="glex"
/locus_tag="NMB0003"
/EC_number="6.1.1.17"
/note="identified by similarity to BGAD:16197; match to protein family HMM PF00749; match to protein family HMM TIGR00464"
/codon_start=1
/transl_table=11
/product="glutamyl-tRNA synthetase"
/protein_id="AAF0482.1"
/db_xref="GI:7225228"
/translation="MTVKTFRPAPSPGYTHIGGVTALFSMAFARHKKSEFLRIEDTDLARSTAESVNIILDGKRWGLNVDNADNVYQTRRPRKVEIAELLEGHAYCYCSKEELPAMREKAKESGATYDRMRPEVGTLPDISDQPVVRFTPLDGVTKMTDLVKGEISIPNALDILLIARADGPTYNFCVVDYDMGVTHVIRGDDHNNPKOINILKADIANLPEYGHLPWILNKGKISKRSGDTVAITDQGMGILPEAMANTYLARLGMAHGDDEFTMEQPIEMFDLKDVSPPSRMDLKKLYWNGEHKITTPGKLAELVPRILA LNDIHTEKPALEADVLELVKDRPDNTLADCEFYFYVQTPPAADVQKMDSEARMLRFAERLGLIEDVNAEAIHDLFPKPCDEBGIMGKLPRLAVCGTAKTSVDVAVLALIGKEVVKRIRA"
complement (2517..3161)
/locus_tag="NMB0004"
complement (2517..3161)
/locus_tag="NMB0004"
/note="identified by similarity to BGAD:132655; match to protein family HMM PF02698"
/codon_start=1
/transl_table=11
/product="BpiH/GdmH-related protein"
/protein_id="AAF0483.1"
/db_xref="GI:7225229"
/translation="NKKRLFCSRNGLRYTLGGFCLSEVPELLVFASVWAVYRTGQVLPYVRADAAVLGAAMDKRSPYRERINHAIALYOSRVRGKIVFGSGTKGYMTABVGRRYALRKGVPARNILPENTSRNTYENINRPLISLTPANGIASVIVSDPHILRAABMAEDGVRYNMSATPTTTPDAKKKIFMLQGVALSLYRLEKMGDRDFPN"
complement (3158..3511)
/gene="arbc"
/locus_tag="NMB0005"
complement (3158..3511)
/gene="arbc"
/locus_tag="NMB0005"
/note="identified by similarity to BGAD:124632; match to protein family HMM PF02698"

Query Match 70.7%; Score 29; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGMAAAACAAATATTCGGTTATAC 41
DB 79111 TATGMAAAACAAATATTCGGTTATAC 79083

RESULT 5
AX043922 349980 bp DNA linear PAT 24-NOV-2000
LOCUS Sequence 1 from Patent WO0066791.
DEFINITION AX043922
ACCESSION AX043922
VERSION AX043922.1 GI:11342850
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.

REFERENCE
1 Pizza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C., Masiagnan, V., Galeotti, C., Mora, M., Ratti, G., Scarselli, M., Scarlato, V., Rappelli, R., Frazer, C. M. and Grandi, G.
Neisseria genomic sequences and methods of their use

TITLE

JOURNAL Patent: WO 0066791-A 1 09-NOV-2000;
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
Source
Location/Qualifiers
1..349980
/organism="Neisseria meningitidis"
/mol_type="unassigned DNA"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces.
seq 1: 1 to 349980 349980 bases
seq 108: 300001 to 649980 349980 bases
seq 109: 600001 to 949980 349980 bases
seq 110: 900001 to 1249980 349980 bases
seq 111: 1200001 to 1549980 349980 bases
seq 112: 1500001 to 1849980 349980 bases
seq 113: 1800001 to 2149980 349980 bases
seq 114: 2100001 to 2272325 172325 bases"

ORIGIN
Query Match 70.7%; Score 29; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGMAAAACAAATATTCGGTTATAC 41
DB 79104 TATGMAAAACAAATATTCGGTTATAC 79076

RESULT 6
NM060146 687 bp DNA linear BCT 11-JUN-1997
LOCUS Neisseria meningitidis CMP-stilic acid synthetase gene, complete
DEFINITION
cgs
U60146.1 GI:1549340
VERSION
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.
1 (bases 1 to 687)
2 (bases 1 to 687)
3 (bases 1 to 687)
4 (bases 1 to 687)
5 (bases 1 to 687)
6 (bases 1 to 687)
7 (bases 1 to 687)
8 (bases 1 to 687)
9 (bases 1 to 687)
10 (bases 1 to 687)
11 (bases 1 to 687)
12 (bases 1 to 687)
13 (bases 1 to 687)
14 (bases 1 to 687)
15 (bases 1 to 687)
16 (bases 1 to 687)
17 (bases 1 to 687)
18 (bases 1 to 687)
19 (bases 1 to 687)
20 (bases 1 to 687)
21 (bases 1 to 687)
22 (bases 1 to 687)
23 (bases 1 to 687)
24 (bases 1 to 687)
25 (bases 1 to 687)
26 (bases 1 to 687)
27 (bases 1 to 687)
28 (bases 1 to 687)
29 (bases 1 to 687)
30 (bases 1 to 687)
31 (bases 1 to 687)
32 (bases 1 to 687)
33 (bases 1 to 687)
34 (bases 1 to 687)
35 (bases 1 to 687)
36 (bases 1 to 687)
37 (bases 1 to 687)
38 (bases 1 to 687)
39 (bases 1 to 687)
40 (bases 1 to 687)
41 (bases 1 to 687)
42 (bases 1 to 687)
43 (bases 1 to 687)
44 (bases 1 to 687)
45 (bases 1 to 687)
46 (bases 1 to 687)
47 (bases 1 to 687)
48 (bases 1 to 687)
49 (bases 1 to 687)
50 (bases 1 to 687)
51 (bases 1 to 687)
52 (bases 1 to 687)
53 (bases 1 to 687)
54 (bases 1 to 687)
55 (bases 1 to 687)
56 (bases 1 to 687)
57 (bases 1 to 687)
58 (bases 1 to 687)
59 (bases 1 to 687)
60 (bases 1 to 687)
61 (bases 1 to 687)
62 (bases 1 to 687)
63 (bases 1 to 687)
64 (bases 1 to 687)
65 (bases 1 to 687)
66 (bases 1 to 687)
67 (bases 1 to 687)
68 (bases 1 to 687)
69 (bases 1 to 687)
70 (bases 1 to 687)
71 (bases 1 to 687)
72 (bases 1 to 687)
73 (bases 1 to 687)
74 (bases 1 to 687)
75 (bases 1 to 687)
76 (bases 1 to 687)
77 (bases 1 to 687)
78 (bases 1 to 687)
79 (bases 1 to 687)
80 (bases 1 to 687)
81 (bases 1 to 687)
82 (bases 1 to 687)
83 (bases 1 to 687)
84 (bases 1 to 687)
85 (bases 1 to 687)
86 (bases 1 to 687)
87 (bases 1 to 687)
88 (bases 1 to 687)
89 (bases 1 to 687)
90 (bases 1 to 687)
91 (bases 1 to 687)
92 (bases 1 to 687)
93 (bases 1 to 687)
94 (bases 1 to 687)
95 (bases 1 to 687)
96 (bases 1 to 687)
97 (bases 1 to 687)
98 (bases 1 to 687)
99 (bases 1 to 687)
100 (bases 1 to 687)
101 (bases 1 to 687)
102 (bases 1 to 687)
103 (bases 1 to 687)
104 (bases 1 to 687)
105 (bases 1 to 687)
106 (bases 1 to 687)
107 (bases 1 to 687)
108 (bases 1 to 687)
109 (bases 1 to 687)
110 (bases 1 to 687)
111 (bases 1 to 687)
112 (bases 1 to 687)
113 (bases 1 to 687)
114 (bases 1 to 687)
115 (bases 1 to 687)
116 (bases 1 to 687)
117 (bases 1 to 687)
118 (bases 1 to 687)
119 (bases 1 to 687)
120 (bases 1 to 687)
121 (bases 1 to 687)
122 (bases 1 to 687)
123 (bases 1 to 687)
124 (bases 1 to 687)
125 (bases 1 to 687)
126 (bases 1 to 687)
127 (bases 1 to 687)
128 (bases 1 to 687)
129 (bases 1 to 687)
130 (bases 1 to 687)
131 (bases 1 to 687)
132 (bases 1 to 687)
133 (bases 1 to 687)
134 (bases 1 to 687)
135 (bases 1 to 687)
136 (bases 1 to 687)
137 (bases 1 to 687)
138 (bases 1 to 687)
139 (bases 1 to 687)
140 (bases 1 to 687)
141 (bases 1 to 687)
142 (bases 1 to 687)
143 (bases 1 to 687)
144 (bases 1 to 687)
145 (bases 1 to 687)
146 (bases 1 to 687)
147 (bases 1 to 687)
148 (bases 1 to 687)
149 (bases 1 to 687)
150 (bases 1 to 687)
151 (bases 1 to 687)
152 (bases 1 to 687)
153 (bases 1 to 687)
154 (bases 1 to 687)
155 (bases 1 to 687)
156 (bases 1 to 687)
157 (bases 1 to 687)
158 (bases 1 to 687)
159 (bases 1 to 687)
160 (bases 1 to 687)
161 (bases 1 to 687)
162 (bases 1 to 687)
163 (bases 1 to 687)
164 (bases 1 to 687)
165 (bases 1 to 687)
166 (bases 1 to 687)
167 (bases 1 to 687)
168 (bases 1 to 687)
169 (bases 1 to 687)
170 (bases 1 to 687)
171 (bases 1 to 687)
172 (bases 1 to 687)
173 (bases 1 to 687)
174 (bases 1 to 687)
175 (bases 1 to 687)
176 (bases 1 to 687)
177 (bases 1 to 687)
178 (bases 1 to 687)
179 (bases 1 to 687)
180 (bases 1 to 687)
181 (bases 1 to 687)
182 (bases 1 to 687)
183 (bases 1 to 687)
184 (bases 1 to 687)
185 (bases 1 to 687)
186 (bases 1 to 687)
187 (bases 1 to 687)
188 (bases 1 to 687)
189 (bases 1 to 687)
190 (bases 1 to 687)
191 (bases 1 to 687)
192 (bases 1 to 687)
193 (bases 1 to 687)
194 (bases 1 to 687)
195 (bases 1 to 687)
196 (bases 1 to 687)
197 (bases 1 to 687)
198 (bases 1 to 687)
199 (bases 1 to 687)
200 (bases 1 to 687)
201 (bases 1 to 687)
202 (bases 1 to 687)
203 (bases 1 to 687)
204 (bases 1 to 687)
205 (bases 1 to 687)
206 (bases 1 to 687)
207 (bases 1 to 687)
208 (bases 1 to 687)
209 (bases 1 to 687)
210 (bases 1 to 687)
211 (bases 1 to 687)
212 (bases 1 to 687)
213 (bases 1 to 687)
214 (bases 1 to 687)
215 (bases 1 to 687)
216 (bases 1 to 687)
217 (bases 1 to 687)
218 (bases 1 to 687)
219 (bases 1 to 687)
220 (bases 1 to 687)
221 (bases 1 to 687)
222 (bases 1 to 687)
223 (bases 1 to 687)
224 (bases 1 to 687)
225 (bases 1 to 687)
226 (bases 1 to 687)
227 (bases 1 to 687)
228 (bases 1 to 687)
229 (bases 1 to 687)
230 (bases 1 to 687)
231 (bases 1 to 687)
232 (bases 1 to 687)
233 (bases 1 to 687)
234 (bases 1 to 687)
235 (bases 1 to 687)
236 (bases 1 to 687)
237 (bases 1 to 687)
238 (bases 1 to 687)
239 (bases 1 to 687)
240 (bases 1 to 687)
241 (bases 1 to 687)
242 (bases 1 to 687)
243 (bases 1 to 687)
244 (bases 1 to 687)
245 (bases 1 to 687)
246 (bases 1 to 687)
247 (bases 1 to 687)
248 (bases 1 to 687)
249 (bases 1 to 687)
250 (bases 1 to 687)
251 (bases 1 to 687)
252 (bases 1 to 687)
253 (bases 1 to 687)
254 (bases 1 to 687)
255 (bases 1 to 687)
256 (bases 1 to 687)
257 (bases 1 to 687)
258 (bases 1 to 687)
259 (bases 1 to 687)
260 (bases 1 to 687)
261 (bases 1 to 687)
262 (bases 1 to 687)
263 (bases 1 to 687)
264 (bases 1 to 687)
265 (bases 1 to 687)
266 (bases 1 to 687)
267 (bases 1 to 687)
268 (bases 1 to 687)
269 (bases 1 to 687)
270 (bases 1 to 687)
271 (bases 1 to 687)
272 (bases 1 to 687)
273 (bases 1 to 687)
274 (bases 1 to 687)
275 (bases 1 to 687)
276 (bases 1 to 687)
277 (bases 1 to 687)
278 (bases 1 to 687)
279 (bases 1 to 687)
280 (bases 1 to 687)
281 (bases 1 to 687)
282 (bases 1 to 687)
283 (bases 1 to 687)
284 (bases 1 to 687)
285 (bases 1 to 687)
286 (bases 1 to 687)
287 (bases 1 to 687)
288 (bases 1 to 687)
289 (bases 1 to 687)
290 (bases 1 to 687)
291 (bases 1 to 687)
292 (bases 1 to 687)
293 (bases 1 to 687)
294 (bases 1 to 687)
295 (bases 1 to 687)
296 (bases 1 to 687)
297 (bases 1 to 687)
298 (bases 1 to 687)
299 (bases 1 to 687)
300 (bases 1 to 687)
301 (bases 1 to 687)
302 (bases 1 to 687)
303 (bases 1 to 687)
304 (bases 1 to 687)
305 (bases 1 to 687)
306 (bases 1 to 687)
307 (bases 1 to 687)
308 (bases 1 to 687)
309 (bases 1 to 687)
310 (bases 1 to 687)
311 (bases 1 to 687)
312 (bases 1 to 687)
313 (bases 1 to 687)
314 (bases 1 to 687)
315 (bases 1 to 687)
316 (bases 1 to 687)
317 (bases 1 to 687)
318 (bases 1 to 687)
319 (bases 1 to 687)
320 (bases 1 to 687)
321 (bases 1 to 687)
322 (bases 1 to 687)
323 (bases 1 to 687)
324 (bases 1 to 687)
325 (bases 1 to 687)
326 (bases 1 to 687)
327 (bases 1 to 687)
328 (bases 1 to 687)
329 (bases 1 to 687)
330 (bases 1 to 687)
331 (bases 1 to 687)
332 (bases 1 to 687)
333 (bases 1 to 687)
334 (bases 1 to 687)
335 (bases 1 to 687)
336 (bases 1 to 687)
337 (bases 1 to 687)
338 (bases 1 to 687)
339 (bases 1 to 687)
340 (bases 1 to 687)
341 (bases 1 to 687)
342 (bases 1 to 687)
343 (bases 1 to 687)
344 (bases 1 to 687)
345 (bases 1 to 687)
346 (bases 1 to 687)
347 (bases 1 to 687)
348 (bases 1 to 687)
349 (bases 1 to 687)
350 (bases 1 to 687)
351 (bases 1 to 687)
352 (bases 1 to 687)
353 (bases 1 to 687)
354 (bases 1 to 687)
355 (bases 1 to 687)
356 (bases 1 to 687)
357 (bases 1 to 687)
358 (bases 1 to 687)
359 (bases 1 to 687)
360 (bases 1 to 687)
361 (bases 1 to 687)
362 (bases 1 to 687)
363 (bases 1 to 687)
364 (bases 1 to 687)
365 (bases 1 to 687)
366 (bases 1 to 687)
367 (bases 1 to 687)
368 (bases 1 to 687)
369 (bases 1 to 687)
370 (bases 1 to 687)
371 (bases 1 to 687)
372 (bases 1 to 687)
373 (bases 1 to 687)
374 (bases 1 to 687)
375 (bases 1 to 687)
376 (bases 1 to 687)
377 (bases 1 to 687)
378 (bases 1 to 687)
379 (bases 1 to 687)
380 (bases 1 to 687)
381 (bases 1 to 687)
382 (bases 1 to 687)
383 (bases 1 to 687)
384 (bases 1 to 687)
385 (bases 1 to 687)
386 (bases 1 to 687)
387 (bases 1 to 687)
388 (bases 1 to 687)
389 (bases 1 to 687)
390 (bases 1 to 687)
391 (bases 1 to 687)
392 (bases 1 to 687)
393 (bases 1 to 687)
394 (bases 1 to 687)
395 (bases 1 to 687)
396 (bases 1 to 687)
397 (bases 1 to 687)
398 (bases 1 to 687)
399 (bases 1 to 687)
400 (bases 1 to 687)
401 (bases 1 to 687)
402 (bases 1 to 687)
403 (bases 1 to 687)
404 (bases 1 to 687)
405 (bases 1 to 687)
406 (bases 1 to 687)
407 (bases 1 to 687)
408 (bases 1 to 687)
409 (bases 1 to 687)
410 (bases 1 to 687)
411 (bases 1 to 687)
412 (bases 1 to 687)
413 (bases 1 to 687)
414 (bases 1 to 687)
415 (bases 1 to 687)
416 (bases 1 to 687)
417 (bases 1 to 687)
418 (bases 1 to 687)
419 (bases 1 to 687)
420 (bases 1 to 687)
421 (bases 1 to 687)
422 (bases 1 to 687)
423 (bases 1 to 687)
424 (bases 1 to 687)
425 (bases 1 to 687)
426 (bases 1 to 687)
427 (bases 1 to 687)
428 (bases 1 to 687)
429 (bases 1 to 687)
430 (bases 1 to 687)
431 (bases 1 to 687)
432 (bases 1 to 687)
433 (bases 1 to 687)
434 (bases 1 to 687)
435 (bases 1 to 687)
436 (bases 1 to 687)
437 (bases 1 to 687)
438 (bases 1 to 687)
439 (bases 1 to 687)
440 (bases 1 to 687)
441 (bases 1 to 687)
442 (bases 1 to 687)
443 (bases 1 to 687)
444 (bases 1 to 687)
445 (bases 1 to 687)
446 (bases 1 to 687)
447 (bases 1 to 687)
448 (bases 1 to 687)
449 (bases 1 to 687)
450 (bases 1 to 687)
451 (bases 1 to 687)
452 (bases 1 to 687)
453 (bases 1 to 687)
454 (bases 1 to 687)
455 (bases 1 to 687)
456 (bases 1 to 687)
457 (bases 1 to 687)
458 (bases 1 to 687)
459 (bases 1 to 687)
460 (bases 1 to 687)
461 (bases 1 to 687)
462 (bases 1 to 687)
463 (bases 1 to 687)
464 (bases 1 to 687)
465 (bases 1 to 687)
466 (bases 1 to 687)
467 (bases 1 to 687)
468 (bases 1 to 687)
469 (bases 1 to 687)
470 (bases 1 to 687)
471 (bases 1 to 687)
472 (bases 1 to 687)
473 (bases 1 to 687)
474 (bases 1 to 687)
475 (bases 1 to 687)
476 (bases 1 to 687)
477 (bases 1 to 687)
478 (bases 1 to 687)
479 (bases 1 to 687)
480 (bases 1 to 687)
481 (bases 1 to 687)
482 (bases 1 to 687)
483 (bases 1 to 687)
484 (bases 1 to 687)
485 (bases 1 to 687)
486 (bases 1 to 687)
487 (bases 1 to 687)
488 (bases 1 to 687)
489 (bases 1 to 687)
490 (bases 1 to 687)
491 (bases 1 to 687)
492 (bases 1 to 687)
493 (bases 1 to 687)
494 (bases 1 to 687)
495 (bases 1 to 687)
496 (bases 1 to 687)
497 (bases 1 to 687)
498 (bases 1 to 687)
499 (bases 1 to 687)
500 (bases 1 to 687)
501 (bases 1 to 687)
502 (bases 1 to 687)
503 (bases 1 to 687)
504 (bases 1 to 687)
505 (bases 1 to 687)
506 (bases 1 to 687)
507 (bases 1 to 687)
508 (bases 1 to 687)
509 (bases 1 to 687)
510 (bases 1 to 687)
511 (bases 1 to 687)
512 (bases 1 to 687)
513 (bases 1 to 687)
514 (bases 1 to 687)
515 (bases 1 to 687)
516 (bases 1 to 687)
517 (bases 1 to 687)
518 (bases 1 to 687)
519 (bases 1 to 687)
520 (bases 1 to 687)
521 (bases 1 to 687)
522 (bases 1 to 687)
523 (bases 1 to 687)
524 (bases 1 to 687)
525 (bases 1 to 687)
526 (bases 1 to 687)
527 (bases 1 to 687)
528 (bases 1 to 687)
529 (bases 1 to 687)
530 (bases 1 to 687)
531 (bases 1 to 687)
532 (bases 1 to 687)
533 (bases 1 to 687)
534 (bases 1 to 687)
535 (bases 1 to 687)
536 (bases 1 to 687)
537 (bases 1 to 687)
538 (bases 1 to 687)
539 (bases 1 to 687)
540 (bases 1 to 687)
541 (bases 1 to 687)
542 (bases 1 to 687)
543 (bases 1 to 687)
544 (bases 1 to 687)
545 (bases 1 to 687)
546 (bases 1 to 687)
547 (bases 1 to 687)
548 (bases 1 to 687)
549 (bases 1 to 687)
550 (bases 1 to 687)
551 (bases 1 to 687)
552 (bases 1 to 687)
553 (bases 1 to 687)
554 (bases 1 to 687)
555 (bases 1 to 687)
556 (bases 1 to 687)
557 (bases 1 to 687)
558 (bases 1 to 687)
559 (bases 1 to 687)
560 (bases 1 to 687)
561 (bases 1 to 687)
562 (bases 1 to 687)
563 (bases 1 to 687)
564 (bases 1 to 687)
565 (bases 1 to 687)
566 (bases 1 to 687)
567 (bases 1 to 687)
568 (bases 1 to 687)
569 (bases 1 to 687)
570 (bases 1 to 687)
571 (bases 1 to 687)
572 (bases 1 to 687)
573 (bases 1 to 687)
574 (bases 1 to 687)
575 (bases 1 to 687)
576 (bases 1 to 687)
577 (bases 1 to 687)
578 (bases 1 to 687)
579 (bases 1 to 687)
580 (bases 1 to 687)
581 (bases 1 to 687)
582 (bases 1 to 687)
583 (bases 1 to 687)
584 (bases 1 to 687)
585 (bases 1 to 687)
586 (bases 1 to 687)
587 (bases 1 to 687)
588 (bases 1 to 687)
589 (bases 1 to 687)
590 (bases 1 to 687)
591 (bases 1 to 687)
592 (bases 1 to 687)
593 (bases 1 to 687)
594 (bases 1 to 687)
595 (bases 1 to 687)
596 (bases 1 to 687)
597 (bases 1 to 687)
598 (bases 1 to 687)
599 (bases 1 to 687)
600 (bases 1 to 687)
601 (bases 1 to 687)
602 (bases 1 to 687)
603 (bases 1 to 687)
604 (bases 1 to 687)
605 (bases 1 to 687)
606 (bases 1 to 687)
607 (bases 1 to 687)
608 (bases 1 to 687)
609 (bases 1 to 687)
610 (bases 1 to 687)
611 (bases 1 to 687)
612 (bases 1 to 687)
613 (bases 1 to 687)
614 (bases 1 to 687)
615 (bases 1 to 687)
616 (bases 1 to 687)
617 (bases 1 to 687)
618 (bases 1 to 687)
619 (bases 1 to 687)
620 (bases 1 to 687)
621 (bases 1 to 687)
622 (bases 1 to 687)
623 (bases 1 to 687)
624 (bases 1 to 687)
625 (bases 1 to 687)
626 (bases 1 to 687)
627 (bases 1 to 687)
628 (bases 1 to 687)
629 (bases 1 to 687)
630 (bases 1 to 687)
631 (bases 1 to 687)
632 (bases 1 to 687)
633 (bases 1 to 687)
634 (bases 1 to 687)
635 (bases 1 to 687)
636 (bases 1 to 687)
637 (bases 1 to 687)
638 (bases 1 to 687)
639 (bases 1 to 687)
640 (bases 1 to 687)
641 (bases 1 to 687)
642 (bases 1 to 687)
643 (bases 1 to 687)
644 (bases 1 to 687)
645 (bases 1 to 687)
646 (bases 1 to 687)
647 (bases 1 to 687)
648 (bases 1 to 687)
649 (bases 1 to 687)
650 (bases 1 to 687)
651 (bases 1 to 687)
652 (bases 1 to 687)
653 (bases 1 to 687)
654 (bases 1 to 687)
655 (bases 1 to 687)
656 (bases 1 to 687)
657 (bases 1 to 687)
658 (bases 1 to 687)
659 (bases 1 to 687)
660 (bases 1 to 687)
661 (bases 1 to 687)
662 (bases 1 to 687)
663 (bases 1 to 687)
664 (bases 1 to 687)
665 (bases 1 to 687)
666 (bases 1 to 687)
667 (bases 1 to 687)
668 (bases 1 to 687)
669 (bases 1 to 687)
670 (bases 1 to 687)
671 (bases 1 to 687)
672 (bases 1 to 687)
673 (bases 1 to 687)
674 (bases 1 to 687)
675 (bases 1 to 687)
676 (bases 1 to 687)
677 (bases 1 to 687)
678 (bases 1 to 687)
679 (bases 1 to 687)
680 (bases 1 to 687)
681 (bases 1 to 687)
682 (bases 1 to 687)
683 (bases 1 to 687)
684 (bases 1 to 687)
685 (bases 1 to 687)
686 (bases 1 to 687)
687 (bases 1 to 687)

FEATURES
source
Location/Qualifiers
1..6

Query Match 68.3%; Score 28; DB 1; Length 687;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

14 ATGGAAAAACAATATTCGGGTATAC 41
 1 ATGGAAAAACAATATTCGGGTATAC 28

RESULT 7
 AC103481 224302 bp DNA linear HTG 10-MAY-2003
 LOCUS Rattus norvegicus clone CH230-20A8, WORKING DRAFT SEQUENCE.
 ACCESSION AC103481.6 GI:30520434
 VERSION HTG; HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.
 KEYWORDS Rattus norvegicus
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Rattus.

REFERENCE
 AUTHORS Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Drepper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gueorgiev, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harrey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyvet, A., Karyathy, S., Kelly, S., Kelly, S., Khan, Z., King, R., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, W., Louie, L., Louie, H., Lozano, R. J., Lu, X., Ma, D., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawney, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milsavajevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Ntali, L., Nwankwelen, O., Okunodu, G., Olampong, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shattman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajda, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, U., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanu, K., Vales, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Wotley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, U., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE
 JOURNAL Unpublished

REFERENCE
 AUTHORS Worley, K. C.
 TITLE Direct Submission
 JOURNAL Submitted (25-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 224302)
 Rattus norvegicus Consortium.

REFERENCE
 AUTHORS Direct Submission
 TITLE Direct Submission
 JOURNAL Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On May 10, 2003 this sequence version replaced gi:25092234.

COMMENT
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence overlap within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GHSS
 Center clone name: CH230-20A8
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 209990 bases at least Q40
 Consensus quality: 212566 bases at least Q30
 Consensus quality: 214461 bases at least Q20
 Estimated insert size: 223285; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 224302: contig of 224302 bp in length.
 location/Qualifiers
 1. 224302
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-20A8"
 /complement(222427..223315)
 /note="clone_boundary
 clone_end:77
 site:BCORI
 end_sequence:BH317689"

ORIGIN
 Query Match 62.9%; Score 25.8; DB 14; Length 224302;
 Best Local Similarity 81.1%; Pred. No. 95;
 Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

2 TTAGGAGTCATATGAAACAAATATTCGGGTATAC 38
 2 TTAGGAGTCATATGAAACAAATATTCGGGTATAC 38

Db 75327 TTGGAGGTCTATGAAACAAATATTCGGGTATAC 75363

RESULT 8
LOCUS AR271714 41 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 32 from patent US 6503744.
ACCESSION AR271714
VERSION AR271714.1 GI:29703259
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
JOURNAL Patent: US 6503744-A 32-07-JAN-2003;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..41
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 7.8e+02;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
1 CTTAGAGGTCATATGAAAAAGTTATTATTCGCTG 35
Db

RESULT 9
LOCUS AR481803 41 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 46 from patent US 6699705.
ACCESSION AR481803
VERSION AR481803.1 GI:47243438
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
JOURNAL Patent: US 6699705-A 46-02-MAR-2004;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..41
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 7.8e+02;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
1 CTTAGAGGTCATATGAAAAAGTTATTATTCGCTG 35
Db

RESULT 10
LOCUS AR527402 41 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 46 from patent US 6723545.
ACCESSION AR527402
VERSION AR527402.1 GI:53914423
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 41)
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Polypeptides having .beta.-1,4-GalNAc transferase activity
JOURNAL Patent: US 6723545-A 46-20-APR-2004;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..41
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 7.8e+02;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
1 CTTAGAGGTCATATGAAAAAGTTATTATTCGCTG 35
Db

RESULT 11
LOCUS AR609681 41 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 46 from patent US 6825019.
ACCESSION AR609681
VERSION AR609681.1 GI:5664981
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Polypeptides having .beta.-1,3-galactosyl transferase activity
JOURNAL Patent: US 6825019-A 46-30-NOV-2004;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..41
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 7.8e+02;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
1 CTTAGAGGTCATATGAAAAAGTTATTATTCGCTG 35
Db

RESULT 12
LOCUS AX934469 41 bp DNA linear PAT 05-JAN-2004
DEFINITION Sequence 46 from Patent WO02074942.
ACCESSION AX934469
VERSION AX934469.1 GI:40641707
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
JOURNAL Patent: WO 02074942-A 46-26-SEP-2002;
National Research Council of Canada (CA)
FEATURES
source Location/Qualifiers
1..41

ORIGIN

/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence:CU-131 5' primer
used to amplify and clone ORF 7a"

Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 7,8e+02;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Cy 1 CTTAGAGGTCATATGCAAAAATAATTCGCG 35
1 CTTAGAGGTCATATGCAAAAATAATTCGCG 35

RESULT 13
LOCUS AE015576 10029 bp DNA linear BCT 02-DEC-2002
DEFINITION Shewanella oneidensis MR-1 section 125 of 457 of the complete
genome

ACCESSION AE015576 AE014299
VERSION AE015576.1 GI:24347020
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS Shewanella oneidensis MR-1
Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
Shewanellaceae; Shewanella.
1 (bases 1 to 10029)
Heidelberg, J., Paulsen, I., Nelson, K., Gaidos, E., Nelson, W.,
Read, T., Eisen, J., Seshadri, R., Ward, N., Meche, B., Clayton, R.,
Meyer, T., Tsapin, A., Scott, J., Beanan, M., Brinkac, L., Daugherty, S.,
Deboy, R., Dodson, R., Durkin, A., Haft, D., Kolonay, J., Madupu, R.,
Peterson, J., Umayam, L., White, O., Wolf, A., Yamaneva, J.,
Weidman, J., Impraim, M., Lee, K., Berry, K., Lee, C., Mueller, J.,
Khouri, H., Gill, J., Uteerback, T., McDonald, L., Feldblyum, T.,
Smith, H., Venter, J., Nealeon, K., and Fraser, C.
Genome sequence of the dissimilatory metal ion-reducing bacterium
Shewanella oneidensis
Nat. Biotechnol. 20 (11), 1118-1123 (2002)
12368813

JOURNAL
PUBMED
REFERENCE
AUTHORS 2 (bases 1 to 10029)
Heidelberg, J., Paulsen, I., Nelson, K., Gaidos, E., Nelson, W.,
Read, T., Eisen, J., Seshadri, R., Ward, N., Meche, B., Clayton, R.,
Meyer, T., Tsapin, A., Scott, J., Beanan, M., Brinkac, L., Daugherty, S.,
Deboy, R., Dodson, R., Durkin, A., Haft, D., Kolonay, J., Madupu, R.,
Peterson, J., Umayam, L., White, O., Wolf, A., Yamaneva, J.,
Weidman, J., Impraim, M., Lee, K., Berry, K., Lee, C., Mueller, J.,
Khouri, H., Gill, J., Uteerback, T., McDonald, L., Feldblyum, T.,
Smith, H., Venter, J., Nealeon, K., and Fraser, C.
Direct Submission
Submitted (12-SEP-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
Location/Qualifiers
1. 10029

TITLE
JOURNAL
FEATURES
source
/organism="Shewanella oneidensis MR-1"
/mol_type="genomic DNA"
/strain="MR-1"
/db_xref="taxon:211586"
complement (93. .584)
/locus_tag="SO1318"
complement (93. .584)
/locus_tag="SO1318"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAN54383.1"
/db_xref="GI:24347021"
/translation="MRILRLVCLIFVPAIPFAHYANGSEMPAEQOLAVKYNALTE
HDVQTSKRYNDSVPEPTARKTGGRFIDPFRAHGVLEVDNFEHNYNAGSL
VVMIGNYHFKGPEQFGPKGIIVAIPEVTSCLKDMKTRHVAEHMDLIDYQTLSDQL
AMQ"

gene
CDS

complement (586. .1212)
/locus_tag="SO1319"
complement (586. .1212)
/locus_tag="SO1319"
/note="similar to GP:6606500; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAN54384.1"
/db_xref="GI:24347022"
/translation="MNNHWYFDFDLCGTAVFALSGALAGRHPDFGVTVASVAVG
GGSIRDALIGATPVFIRDPYIVIVATVACILVRRPKMEYVLPVADALAL
FTVIGAKALMGISGMIAVVGILITGVGGIIRIDLCROIPLVLRTEIVTASIIIG
IGTVSLACGMGERTALPLAMASALIRLCAIKHLSIPADLTKE"

gene
CDS

complement (1209. .1598)
/locus_tag="SO1320"
complement (1209. .1598)
/locus_tag="SO1320"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAN54385.1"
/db_xref="GI:24347023"
/translation="MFENMYSLTTPALIPPAISLILLYATNRPFAALAIROISGD
KPVHQDKNSQSRIRIRPMQEGASFPALCVCMFTIYIGAKTGSITFGASILL
LYSLVSTETRSVDALNHIKEMSK"
complement (1642. .2631)
/locus_tag="SO1321"
complement (1642. .2631)
/locus_tag="SO1321"
/note="similar to GB:M73980, SP:P46531, and PID:338675;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cobd-related protein"
/protein_id="AAN54386.1"
/db_xref="GI:24347024"
/translation="MHSEFYQOLVEIGALPEFGLVLPFALLIARLAPREMOPLIV
FSLAKQIAKAVNRPGRAPSOQATRGFLAMLLVPPRAITTLLELAPFWPREFV
LYLCLSDGFSQVADDEIAKALNRBDNSAKLLQPVVVRDITLLELAPFWPREFV
TAPVYGSTSTIFPAIAGAPVAVLVRMKQELSWPPIQPKYQYFCRVHTLSALIA
IPAMVMSIAIIGCPKALALLFTSKNHPALRGVMSVLVANIIRLEIGSPQKFSG
GRIVAKVVGPOHHEMIAPAIKLTSAICAIWSPFTLPLFIWAGRWLQTL"
complement (2663. .3373)
/gene="pfs"
/locus_tag="SO1322"
complement (2663. .3373)
/gene="pfs"
/locus_tag="SO1322"
/note="similar to GB:X70848, GB:J04759, GB:M63960,
SP:P08129, SP:P36873, SP:P37140, PID:190281, PID:190516,
PID:35451, PID:484316, PID:506772, and PID:531476;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="5-methylthioadenosine
nucleosidase/S-adenosylhomocysteine nucleosidase"
/protein_id="AAN54387.1"
/db_xref="GI:24347025"
/translation="MKIGIIGAMEBEVAHLIAAMTNTSSQRTADIETAGTLAKDQV
VTRSGIGVAAISIAITLLIKYAPDAVINTGSAQGFVDSLIGDIVISSRYRHVDV
TAFQIGCMQAQPAAPFIPAVIVEAANKAIAQGEVAIRIGLICGDSFICDPVRTK
TMLHFFPMAACBEGAAIAOYCHQFQVFPVVISLSLDNANDSVVDYATIVKAGYH
SALMWMLLEQLQLEPNSVK"
3372. .3578
/locus_tag="SO1323"
3372. .3578
/locus_tag="SO1323"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=11

gene
CDS

complement (1642. .2631)
/locus_tag="SO1321"
complement (1642. .2631)
/locus_tag="SO1321"
/note="similar to GB:M73980, SP:P46531, and PID:338675;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cobd-related protein"
/protein_id="AAN54386.1"
/db_xref="GI:24347024"
/translation="MHSEFYQOLVEIGALPEFGLVLPFALLIARLAPREMOPLIV
FSLAKQIAKAVNRPGRAPSOQATRGFLAMLLVPPRAITTLLELAPFWPREFV
LYLCLSDGFSQVADDEIAKALNRBDNSAKLLQPVVVRDITLLELAPFWPREFV
TAPVYGSTSTIFPAIAGAPVAVLVRMKQELSWPPIQPKYQYFCRVHTLSALIA
IPAMVMSIAIIGCPKALALLFTSKNHPALRGVMSVLVANIIRLEIGSPQKFSG
GRIVAKVVGPOHHEMIAPAIKLTSAICAIWSPFTLPLFIWAGRWLQTL"
complement (2663. .3373)
/gene="pfs"
/locus_tag="SO1322"
complement (2663. .3373)
/gene="pfs"
/locus_tag="SO1322"
/note="similar to GB:X70848, GB:J04759, GB:M63960,
SP:P08129, SP:P36873, SP:P37140, PID:190281, PID:190516,
PID:35451, PID:484316, PID:506772, and PID:531476;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="5-methylthioadenosine
nucleosidase/S-adenosylhomocysteine nucleosidase"
/protein_id="AAN54387.1"
/db_xref="GI:24347025"
/translation="MKIGIIGAMEBEVAHLIAAMTNTSSQRTADIETAGTLAKDQV
VTRSGIGVAAISIAITLLIKYAPDAVINTGSAQGFVDSLIGDIVISSRYRHVDV
TAFQIGCMQAQPAAPFIPAVIVEAANKAIAQGEVAIRIGLICGDSFICDPVRTK
TMLHFFPMAACBEGAAIAOYCHQFQVFPVVISLSLDNANDSVVDYATIVKAGYH
SALMWMLLEQLQLEPNSVK"
3372. .3578
/locus_tag="SO1323"
3372. .3578
/locus_tag="SO1323"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=11

gene
CDS

complement (1642. .2631)
/locus_tag="SO1321"
complement (1642. .2631)
/locus_tag="SO1321"
/note="similar to GB:M73980, SP:P46531, and PID:338675;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cobd-related protein"
/protein_id="AAN54386.1"
/db_xref="GI:24347024"
/translation="MHSEFYQOLVEIGALPEFGLVLPFALLIARLAPREMOPLIV
FSLAKQIAKAVNRPGRAPSOQATRGFLAMLLVPPRAITTLLELAPFWPREFV
LYLCLSDGFSQVADDEIAKALNRBDNSAKLLQPVVVRDITLLELAPFWPREFV
TAPVYGSTSTIFPAIAGAPVAVLVRMKQELSWPPIQPKYQYFCRVHTLSALIA
IPAMVMSIAIIGCPKALALLFTSKNHPALRGVMSVLVANIIRLEIGSPQKFSG
GRIVAKVVGPOHHEMIAPAIKLTSAICAIWSPFTLPLFIWAGRWLQTL"
complement (2663. .3373)
/gene="pfs"
/locus_tag="SO1322"
complement (2663. .3373)
/gene="pfs"
/locus_tag="SO1322"
/note="similar to GB:X70848, GB:J04759, GB:M63960,
SP:P08129, SP:P36873, SP:P37140, PID:190281, PID:190516,
PID:35451, PID:484316, PID:506772, and PID:531476;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="5-methylthioadenosine
nucleosidase/S-adenosylhomocysteine nucleosidase"
/protein_id="AAN54387.1"
/db_xref="GI:24347025"
/translation="MKIGIIGAMEBEVAHLIAAMTNTSSQRTADIETAGTLAKDQV
VTRSGIGVAAISIAITLLIKYAPDAVINTGSAQGFVDSLIGDIVISSRYRHVDV
TAFQIGCMQAQPAAPFIPAVIVEAANKAIAQGEVAIRIGLICGDSFICDPVRTK
TMLHFFPMAACBEGAAIAOYCHQFQVFPVVISLSLDNANDSVVDYATIVKAGYH
SALMWMLLEQLQLEPNSVK"
3372. .3578
/locus_tag="SO1323"
3372. .3578
/locus_tag="SO1323"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=11

gene
CDS

complement (1642. .2631)
/locus_tag="SO1321"
complement (1642. .2631)
/locus_tag="SO1321"
/note="similar to GB:M73980, SP:P46531, and PID:338675;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cobd-related protein"
/protein_id="AAN54386.1"
/db_xref="GI:24347024"
/translation="MHSEFYQOLVEIGALPEFGLVLPFALLIARLAPREMOPLIV
FSLAKQIAKAVNRPGRAPSOQATRGFLAMLLVPPRAITTLLELAPFWPREFV
LYLCLSDGFSQVADDEIAKALNRBDNSAKLLQPVVVRDITLLELAPFWPREFV
TAPVYGSTSTIFPAIAGAPVAVLVRMKQELSWPPIQPKYQYFCRVHTLSALIA
IPAMVMSIAIIGCPKALALLFTSKNHPALRGVMSVLVANIIRLEIGSPQKFSG
GRIVAKVVGPOHHEMIAPAIKLTSAICAIWSPFTLPLFIWAGRWLQTL"
complement (2663. .3373)
/gene="pfs"
/locus_tag="SO1322"
complement (2663. .3373)
/gene="pfs"
/locus_tag="SO1322"
/note="similar to GB:X70848, GB:J04759, GB:M63960,
SP:P08129, SP:P36873, SP:P37140, PID:190281, PID:190516,
PID:35451, PID:484316, PID:506772, and PID:531476;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="5-methylthioadenosine
nucleosidase/S-adenosylhomocysteine nucleosidase"
/protein_id="AAN54387.1"
/db_xref="GI:24347025"
/translation="MKIGIIGAMEBEVAHLIAAMTNTSSQRTADIETAGTLAKDQV
VTRSGIGVAAISIAITLLIKYAPDAVINTGSAQGFVDSLIGDIVISSRYRHVDV
TAFQIGCMQAQPAAPFIPAVIVEAANKAIAQGEVAIRIGLICGDSFICDPVRTK
TMLHFFPMAACBEGAAIAOYCHQFQVFPVVISLSLDNANDSVVDYATIVKAGYH
SALMWMLLEQLQLEPNSVK"
3372. .3578
/locus_tag="SO1323"
3372. .3578
/locus_tag="SO1323"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=11

```
/product="hypothetical protein"
/protein_id="AA054388.1"
/db_xref="GI:24347026"
/translation="MGDPDPTSCVPLMMMDGMAANIPOLCFVTVGLRLDRDNT
KQGLGNALIKRAPIYSFVISHL"
complement(3641..5047)
/gene="gltD"
/locus_tag="SO1324"
/complement(3641..5047)
/gene="gltD"
/locus_tag="SO1324"
/notice="similar to SP:P09832; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="glutamate synthase, small subunit"
/protein_id="AA054389.1"
/db_xref="GI:24347027"
/translation="MSNDPQFLTEVGRKDPTEKPAARSTQPIETIYQPAQVSEQAD
RCLDGNPYCEWKCPLAHYIPMLKLAQOGRIMEADLVHENTNPEIGRCPODRL
CGCTLNADPGATVIGVETKYITDIAIGRPMPSKVPKREKVAIVGAPAGIGC
ADILANGKAVPDKYPOIGELITGIPAFKIDKAVATRTVLEMGTEKLGTV
GQDAFNOULEBIDAVFLGMYTAKKALBESDAQVQALPYLIGNTHLMGSTP
DMPYINLAKRQVAVIGGQDPAVCVATVROQASVTCAYRDEANMPGSRREVQAR
BEGVNFLENROQVAKTQDQKTVGECVETQKQKADASGRQAEAVIGSEOLLADAV
IAPGROPSPAPMPADYGIELDQMGVSKAQADNPFTTNKVFAGGDMVAGSLVV
TAIARGDRAAQOILNVLID"
complement(5054..9502)
/gene="gltB"
/locus_tag="SO1325"
/complement(5054..9502)
/gene="gltB"
/locus_tag="SO1325"
/notice="identified by match to PPM protein family HMM
PF00310"
/codon_start=1
/transl_table=11
/product="glutamate synthase, large subunit"
/protein_id="AA054390.1"
/db_xref="GI:24347028"
/translation="MSLYHPSFERDNCGFLIAQNDGEASHRIYRTAIGLDRMKRG
GIASDRTGDCGGLMOLPLQPEFAVAENMDHLRKPVGMPLFQDEBELADQKFI
LERELRETLISAGRWKVPVNDVLSIGKSSLPQIYQVLIAPIGMREKDLERLYM
ARRLEOQITDDKDFYVASLGGVIVYKGLMMPADLPAPYPLADIRLKSSITCFHOR
PSTNTPKMPPLAOPPRVLAHNEITITGNRMAARAKRNSPLIPLOQAAPVNE
TGSBSSIDNMLBMLSGMDIRAMRLIIPAWGNSNPEMDLAFYDPNSMEHPW
DGPAGIVTINGHAAQAVDRNGLRSPRYVITGDRILITLASEGMDVADYETEKGRV
GPELIVLDTLNGRLYOSFEIINDLKRHPYKEMMAKNSRLIPALQTEHQHSEL
SPBOLYOQOKGYRRELEQVITVLAHQGEAETGSMGDDPMAVLSKRSRLYVFR
QKQAVTNPPLDPLREKVMGLATCIGRQNFETTHAYRVWNSPILPSPDNOL
IGLSTYTRANI VLTNTERKGLBQATIRITSEARLARSGLTLLISDRAIDKSAQV
IPVANAQAQOOLVSKSLCDNTNII VETASRDHFAVLGFGATLAIPYLVESI
SDLAQRDLTDTALMNFYRQIEKGLIKMSKMSIISTGSRCSQCFEALGIGSDVI
ELCFKGVISRIEGASPEDIANQDLHAAVRAHLPQGLIKYDGEAEHCFEPDV
VNTLOASLIDKNFATYKKKPAELVDRNPATLRDLGIRAGOTAIELANNVBAASNYPR
PDSAMSTGALSPEAHALATAMNLDGRSNSGEGEDAPRPNTERNSAIQIASAR
GVTHAYVYNADVLOIKVAQAKPQEGQPEKRSVETAGLHARPGVTLISPEPHD
IYISIDLAQOLFDLQINTKALISVLTSEBQVGIATGAAVADMILTISGYDGTG
ASPIYVSKAGSPWELGLAEVHOSLVENGRLHRIKLIQVDDGLKTGTDVITKALGAE
FGFTVPMIATGCKTARI CHLNNCATGATODKLDNDHYGLPERWNTYEFVAVEN
REMAATLGVSKPEDVGRSEWMLTLEGGTDGROGLDAPILYOPNVRASRTWQETN
PPADILNHLNCCOASAVDKGSPDAYSINNTDRVGARLSYIATTTGYSKSA
PIKSPKSAQSGRFGWNSPGIELKCGDANDYVKGMSGGLVYPLVPSPPQERS
AIVGNTCYGATNGRFFAQAQGERFAVRNSGALAVGSLDNGCEWTSQIVVVLGK
TGVMFGAQTGFAVVDQPGFNRNVSSELDVTDKLESPLOOQHLKGLIREHVAETG
SEHAMMILSDPAMWDCFVLVPEKNIAVAIDLKLEQSSPELAVKAG"
```

ORIGIN

```
Query Match 61.5%; Score 25.2; DB 1; Length 10029;
Best Local Similarity 78.9%; Pred. No. 2.8e+02;
Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
3 TAGAGGTCAATGTGAAAAACAAATATTCGCGTTATA 40
```

```
Db 388 TTGGCGGTCTTATCGAAAAAACACTATCGCGTTATA 425
RESULT 14
AE015943/c 301439 bp DNA linear BCT 06-FEB-2003
LOCUS AE015943
DEFINITION Clostridium tetani E88, section 8 of 10 of the complete genome.
ACCESSION AE015943
VERSION AE015943.1
KEYWORDS GI:28204047
SOURCE Clostridium tetani E88
ORGANISM Clostridium tetani E88
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
1 (bases 1 to 301439)
Bruggemann, H., Baumer, S., Fricke, W.F., Wierer, A., Liesegang, H.,
Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A. and
Gottschalk, G.
The genome sequence of Clostridium tetani, the causative agent of
tetanus disease
Proc. Natl. Acad. Sci. U.S.A. 100 (3), 1316-1321 (2003)
2 (bases 1 to 301439)
Bruggemann, H., Baumer, S., Fricke, W.F., Wierer, A., Liesegang, H.,
Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A. and
Gottschalk, G.
Direct Submission
Submitted (31-OCT-2002) Goettingen Genomics Laboratory, Institute
of Microbiology and Genetics, Georg-August University,
Grisebachstr. 8, Goettingen 37077, Germany
Location/Qualifiers
1..301439
/organism="Clostridium tetani E88"
/mol_type="Genomic DNA"
/strain="Massachusetts"
/sub_strain="E88"
/db_xref="taxon:212717"
92..778
/locus_tag="CTC01981"
92..778
/locus_tag="CTC01981"
/codon_start=1
/transl_table=11
/product="L-serine dehydratase beta subunit"
/protein_id="AA036487.1"
/db_xref="GI:28204048"
779..1651
/locus_tag="CTC01982"
779..1651
/locus_tag="CTC01982"
/codon_start=1
/transl_table=11
/product="L-serine dehydratase alpha subunit"
/protein_id="AA036488.1"
/db_xref="GI:28204049"
/translation="MFVNTGELIDICKENKLIKWEGLNVELETGETKDFLIDKN
KTLVMOQSANLGLKEKYSVGLIGDVAVGLANTYSBDKNTLTSGLYKAAARALSS
ENVAMKQIVAPTAGSGGILIPALISAGELANSDEDLVALPTASVGVYIAKNAT
LAGAGGCAQCGSAAAMAASIVELMGSGTVDGSLDAAYIVIKNIIGLVCPVGLIVE
IPCKRNAAGVVSALTSDADMAVSKAIPDDIVLYAMKYGRQLPHELRATALGGLA
TTPAGIKTKKVER"
```

gene
CDS
/product="response regulator protein"
/protein_id="AA036489.1"
/db_xref="GI:28204050"
/translation="MVNWKLSYMELWLMKEVAKAILDNVSESNITDVRILIDPIA
KRIIVCYKREBKLEKSCSPFKGKESCDNCISMRATINESKTPVAKIYESKDNKIYMT
SPMIHGNNKAFVIELVKDIMGTGIIIDEVENATECEIYDIIEKNRRILIKDPLEVYNR
LPIKRLRPVDELAKNANIPLSIMIIDIVYKEINDYTGHAIGFVLKQSKOECRELL
KEYDSMWRSGDEFLIALRDVDRKVLKIKETIISKTRWELTYKNSCIKMSISGCI
YIYESENIDYVELIDRAKNNYNAKEMGRDNIYVN"
2788..3168
/gene="gntR"
/locus_tag="CTC01984"
2788..3168
/gene="gntR"
/locus_tag="CTC01984"
/codon_start=1
/transl_table=11
/product="transcriptional regulator, gntR family"
/protein_id="AA036490.1"
/db_xref="GI:28204051"
/translation="MINIDSSKSIYQOIVDOIKENILKRLKPGKLPVSRMSLSL
LTVNNATYSKAYOELEROKTETIRGRGTIYSRYKPKPMDDDKLFIKEIKKLIIEA
LYMGVKKKFINIDELKYKDFRRM"
3173..4069
/locus_tag="CTC01985"
3173..4069
/locus_tag="CTC01985"
/codon_start=1
/transl_table=11
/product="ABC transporter ATP-binding protein"
/protein_id="AA036491.1"
/db_xref="GI:28204052"
/translation="MLEISNVKKLGNLLINNVNLVSKSIFGLIGENGAGKTYLLI
KQITGIYKDSGNIKIPNEVEENTKIKELGIYADENQYIPSPVSEMIKFPGLTYKVK
KSMRPNELNKTPIPIDRIKELSKMKRLSIGMLNLSIMPELIVLDEPTSLDLYV
IKKKIINILIDVAKNNNTPIITSSHHIGDLRIQDNVAFIKGRTIYVNNIDPMKNNII
KKLQVFPKNNPKIQSDDEVILTISIGRVNVITTKFSKLEKELKSMASLPTEILD
LSLEMFVYVSEDEHNEKILIA"
4041..5114
/locus_tag="CTC01986"
4041..5114
/locus_tag="CTC01986"
/codon_start=1
/transl_table=11
/product="membrane spanning protein"
/protein_id="AA036492.1"
/db_xref="GI:28204053"
/translation="NNTWKKFLHNPLFKDFKSARWISLLVLLIKPKLLIPSPILN
TIKSNLNTDQISPGNMTPKRAEMAYWRKNNLAYSDDMIYITLIVYLLILFKRNR
NNSTYSFPRMSPFKREBEIYKIMLNGMSISIPPTITLISIPKNTMYNDPIYII
VQWITLNLITLAFSTFPIQTMGNQITIASIIGALVIGIPGCIWMSIDPLFSLHY
KNPIIKPIDIYTELKVMYIWMSPDYISIQKGEYTRVSYKNFQIKIILILLCIVAY
VALTEGFLASVYITDKCIKKYYSK"
5132..5824
/locus_tag="CTC01987"
5132..5824
/locus_tag="CTC01987"
/codon_start=1
/transl_table=11
/product="transporter"
/protein_id="AA036493.1"
/db_xref="GI:28204054"
/translation="MDYITLCENDLYKKYGNKIYVNNLNLKIPKKGITGLLPGNCGKT
TILKIVYALNNTSGQITIDGKSPGVDTKRIYAYIPDEFLYQWMSIAEALNPKFKESE
EEDPVKASAMNTNLNDSKOKIKSLSGMQEKNIISLTFSRKSLFLIDPLAAVYVY
VRDKIIEIILNPFKNSIILISTHILINDIYELFNEVVFINDGKTIILGDVLANKNKKYK
KKSIDLPKEKII"
5825..6607
/locus_tag="CTC01988"
5825..6607
/locus_tag="CTC01988"
/codon_start=1

[illegible]

Search completed: April 7, 2006, 18:04:55
 Job time : 285.271 secs

DEFINITION Human DNA sequence from clone RP11-44D23 on chromosome 10, complete sequence.

ACCESSION

AL589943

VERSION

AL589943.26 GI:20068452

KEYWORDS

HTG. Homo sapiens (human)

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens (human)

REFERENCE

1 (bases 1 to 139325)

AUTHORS

Pelam, S.

TITLE

Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk

JOURNAL

On Apr 7, 2002 this sequence version replaced gi:19848045.

COMMENT

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

FEATURES

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

FEATURES

http://www.sanger.ac.uk/projects/C.elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at

FEATURES

http://www.sanger.ac.uk/HGP/Chr10

FEATURES

RP11-44D23 is from the library RPCT-11.1 constructed by the group of Pieter de Jong. For further details see

FEATURES

http://www.chori.org/bacpac/home.htm

FEATURES

VECTOR: pBACe3.6

FEATURES

----- Genome Center

FEATURES

Center: Wellcome Trust Sanger Institute

FEATURES

Center code: SC

FEATURES

Web site: http://www.sanger.ac.uk

FEATURES

Contact: vegas@sanger.ac.uk

FEATURES

FEATURES

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

Location/Qualifiers

FEATURES

1..139325

FEATURES

/organism="Homo sapiens"

FEATURES

/mol_type="genomic DNA"

FEATURES

/db_xref="taxon:9606"

FEATURES

/chromosome="10"

FEATURES

/clone_1ib="RPCT-11.1"

FEATURES

2000

FEATURES

/note="Clone_right_end: RP11-49L2"

FEATURES

6873

FEATURES

/note="Tandem repeat. Forced join. Gap size estimated to be approximately 170bp by restriction digest data."

FEATURES

71187..71380

FEATURES

/note="Single clone region. Assembly confirmed by restriction digest data."

FEATURES

117145

FEATURES

/note="Clone_left_end: RP11-265G8"

FEATURES

139325

FEATURES

/note="Clone_right_end: RP11-44D23"

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

Query Match 60.5%; Score 24.8; DB 8; Length 139325;

Best Local Similarity 80.6%; Pred. No. 2.2e+02;

Matches 29; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

4 AGAGGTCATATGGAACAAATATTCGGTTAT 39

118620 AGGATGCATATGGAACAAATATTCGGTTAT 118585

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr 1 7, 2006, 15:58:06 ; Search time 46.2484 Seconds
(without alignments)
5908.370 Million cell updates/sec

Title: US-09-211-691-3

Perfect score: 41

Sequence: 1 cttagagagctcatatcgtaa.....acaatatcgcgtrtacc 41

Scoring table: IDENTITY_NTC

Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

N_Geneseq_21.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	41	2	AAx84282
2	29	70.7	65632	3	AAa81502
3	29	70.7	110000	3	AAa81490 N. mening
4	29	70.7	349980	3	AAa81490 N. mening
5	28	68.3	687	3	AAa224472 N. mening
6	25.4	62.0	41	3	AAa53735 Primer fo
7	25.4	62.0	41	3	AAa53735 Primer fo
8	24.8	60.5	165199	6	ABR83460 Human CDN
9	24.4	59.5	4356	12	AD020395
10	24.4	59.5	7183	11	ACN90426
11	23	56.1	444	3	AAc97275
12	23	56.1	1024	6	ABx65592 Helicobac
13	23	56.1	1680	8	ACA344620 Prokaryot
14	23	56.1	1780	2	AAx14256 H. pylori
15	23	56.1	4391	14	AAa61394 Bacillus
16	22.6	55.1	1538	6	ABQ70554
17	22.6	55.1	1628	3	AAc47551 Arabidops
18	22.6	55.1	1633	3	AAc35365 Arabidops
19	22.6	55.1	2000	8	ADA69010 Arabidops

C 20	22.6	55.1	14012	8	AAU50004	Aad50004 Human Ras
C 21	22.6	55.1	110000	6	ABA03041_19	Continuation (20 C
C 22	22.4	54.6	542	4	AAK62272	Aak62272 Human imm
C 23	22.4	54.6	1963	3	AAK83350	Aak83350 DNA encod
C 24	22.4	54.6	4121	6	ABL59256	AbL59256 Adenosine
C 25	22.4	54.6	4954	4	AAK82088	Aak82088 Human imm
C 26	22.2	54.1	3089	4	ABL26828	AbL26828 Drosophi
C 27	22.2	54.1	4836	6	ABL32111	AbL32111 Human imm
C 28	22	53.7	31	3	AAZ24470	Aaz24470 N. mening
C 29	22	53.7	62	2	AAV81717	Aav81717 Neisseria
C 30	22	53.7	690	8	ACA29466	ACA29466 Prokaryot
C 31	22	53.7	1263	3	ACA28653	ACA28653 Prokaryot
C 32	22	53.7	1401	9	ADA32829	Ada32829 DNA encod
C 33	22	53.7	1422	10	ADC92660	Adc92660 B. faeciu
C 34	21.8	53.2	550	5	ABV56670	Abv56670 Human pro
C 35	21.8	53.2	1734	13	ADT46561	Adt46561 Bacterial
C 36	21.8	53.2	6075	2	AAQ27451	Aaq27451 Type A hu
C 37	21.8	53.2	25944	11	ADL27125	ADL27125 Mouse gen
C 38	21.8	53.2	26317	9	ADA66331	Ada66331 Mouse IgG
C 39	21.8	53.2	26320	9	ADA03047	Ada03047 Mouse IgG
C 40	21.8	53.2	26320	10	ADB72785	AdB72785 Mouse IgG
C 41	21.8	53.2	74828	12	ADQ97992	Adq97992 Human can
C 42	21.8	53.2	110000	14	ADZ13747_1	Continuation (2 of
C 43	21.8	53.2	347814	12	ADQ59440	Adq59440 Human can
C 44	21.4	52.2	1020	10	ACF68472	AcF68472 Photorhab
C 45	21.4	52.2	2838	10	ADe99742	ADe99742 Bacterial

ALIGNMENTS

RESULT 1
ID AAX84282 standard; DNA; 41 BP.
XX AAX84282;
AC XX
XX XX
DT 08-SRP-1999 (first entry)
XX XX
DE PCR primer for Neisseria CMP-Neu5Ac synthetase coding sequence.
XX XX
XX Beta-1,4-galactosyltransferase; IgCB; fusion protein; catalytic domain;
XX glycosyltransferase; accessory enzyme; nucleotide sugar formation;
XX saccharide donor; oligosaccharide synthesis; CMP-Neu5Ac synthetase;
XX carbohydrate structure development; PCR primer; ss.
XX XX
OS Synthetic.
OS Neisseria sp.
XX XX
PN WO9931224-A2.
XX XX
PD 24-JUN-1999.
XX XX
PF 15-DEC-1998; 98WO-CA001180.
XX XX
PR 15-DEC-1997; 97US-0069443P.
PR 14-DEC-1998; 98US-00211691.
XX XX
PA (CANADA) NAT RES COUNCIL CANADA.
PI Gilbert M, Young NM, Wakarchuk WM;
XX WPI, 1999-395174/33.
XX XX
PT A new glycosyltransferase fusion protein useful in the enzymatic
XX synthesis of oligosaccharides.
XX XX
PS Example 1; Page 40; 63pp; English.
CC This sequence represents a PCR primer for the Neisseria CMP-Neu5Ac
CC synthetase coding sequence. The invention relates to a nucleic acid
CC encoding a fusion protein that comprises a glycosyltransferase catalytic
CC domain and a catalytic domain from an accessory enzyme that is involved

in formation of a nucleotide sugar which is a saccharide donor for a glycosyltransferase reaction. The fusion protein is useful in the enzymatic synthesis of oligosaccharides. The fusion proteins are able to catalyze more than one reaction involved in the enzymatic synthesis. This is useful for the development of therapeutic agents that have specific carbohydrate structures. Carbohydrates are involved in recognition elements on the surface of cells. The fusion protein can be used for the synthesis of both natural carbohydrates and synthetic derivatives with novel properties. The fusion polypeptide allows two glycosyltransferase reactions in a single vessel, provides improved yields of end products. Additionally, cleanup and disposal of extra solvents and by-products is reduced. The fusion protein can also use directly different donor analogues and various acceptors with a terminal galactose residue

Sequence 41 BP; 16 A; 5 C; 9 G; 11 T; 0 U; 0 Other;

Query Match 100.0%; Score 41; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAAATAATTTGCGTTATAC 41
DB 1 CTTAGAGGTCATATGAAAAAATAATTTGCGTTATAC 41

RESULT 2
AAA81502
ID AAA81502 standard; DNA; 65632 BP.

AC AAA81502;
XX
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gmm_49 SEQ ID NO:49.

XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KM Meningococcus B; MenB; ds.
XX
XX Neisseria meningitidis.
OS
XX WO200022430-A2.
XX
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-US023573.
XX
XX 09-OCT-1998; 98US-0103794P.
XX 30-APR-1999; 99US-0132068P.
XX
XX (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC,
PI Massignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V,
PI Rappelli R, Pizsa M;
XX
XX WPI; 2000-318079/27.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be used
XX in the diagnosis and treatment of N. meningitidis infection and other
XX Neisseria infections, for example, N. gonorrhoea.
XX
XX Claim 7; Page 1331-1350; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic proteins
XX from Neisseria genomic sequences. AAA81453 to AAA82414 represent
XX specifically claimed Neisseria meningitidis genomic DNA sequences;
XX AAA81260 to AAA81303 and AAA83620 to AAA82563 represent Neisseria DNA
XX sequences and their corresponding proteins; AAA81254 to AAA81259 and
XX AAA81304 to AAA81321 represent PCR primers used in the isolation of
XX Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent
XX Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
XX used in the exemplification of the present invention. The nucleic acid

sequences, protein sequences, and antibodies against them, can be used in the manufacture of a composition. The composition can be used as a medicament (or in the manufacture of a medicament) for treating, preventing or diagnosing infection due to Neisseria bacteria. For example, some of the identified proteins could be components of vaccines against Meningococcus B, against all serotypes; and/or against all pathogenic Neisseriae. Identification of sequences from the bacterium will also facilitate production of biological probes, particularly organism-specific probes. Attempts to make efficacious Meningococcus B vaccines have failed mainly due to antigen tolerance. Multivalent CC vaccines have also been tried but none have successfully overcome CC antigenic variability. The provision of further, complete sequences may CC provide an opportunity to identify secreted or surface exposed proteins CC that may be presumed targets for the immune system and which are not CC antigenically variable or at least more conserved than other more CC variable regions

Sequence 65632 BP; 16704 A; 16327 C; 15474 G; 17126 T; 0 U; 1 Other;

Query Match 70.7%; Score 29; DB 3; Length 65632;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGAAAAACAAATATTTGCGTTATAC 41
DB 55054 TATGAAAAACAAATATTTGCGTTATAC 55082

RESULT 3
AAA81490_00/C
WP Sequence split into 15 fragments LOCUS AAA81490 Accession Aaa81490
WP Fragment Name Begin End
WP AAA81490_00 1 110000
WP AAA81490_01 100001 210000
WP AAA81490_02 200001 310000
WP AAA81490_03 300001 410000
WP AAA81490_04 400001 510000
WP AAA81490_05 500001 610000
WP AAA81490_06 600001 710000
WP AAA81490_07 700001 810000
WP AAA81490_08 800001 910000
WP AAA81490_09 900001 1010000
WP AAA81490_10 1000001 1110000
WP AAA81490_11 1100001 1210000
WP AAA81490_12 1200001 1310000
WP AAA81490_13 1300001 1410000
WP AAA81490_14 1400001 1437668
ID AAA81490 standard; DNA; 1437668 BP.
XX
XX AAA81490;
XX
XX 04-DEC-2000 (first entry)
XX
XX N. meningitidis B full length genome DNA sequence SEQ ID NO:1068.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KM Meningococcus B; MenB; ds.
XX
XX Neisseria meningitidis.
OS
XX WO200022430-A2.
XX
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-US023573.
XX
XX 09-OCT-1998; 98US-0103794P.
XX 30-APR-1999; 99US-0132068P.
XX
XX (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;

PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V;
 PI Rappoli R, Pizsa M;
 XX
 DR WPI; 2000-318079/27.

PT Isolated nucleotide sequences of *Neisseria meningitidis* which can be used
 PT in the diagnosis and treatment of *N. meningitidis* infection and other
 PT *Neisseria* infections, for example, *N. gonorrhoea*.
 XX

PS Claim 7; Page 866-1272; 1760pp; English.

XX
 CC The present invention describes methods of obtaining immunogenic proteins
 CC from *Neisseria* genomic sequences. AAA81453 to AAA82414 represent
 CC specifically claimed *Neisseria meningitidis* genomic DNA sequences;
 CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent *Neisseria* DNA
 CC sequences and their corresponding proteins; AAA81254 to AAA81259 and
 CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
 CC *Neisseria meningitidis* DNA sequences; and AAA81322 to AAA81452 represent
 CC *Neisseria meningitidis* MsmB polynucleotide ORF sequences, which are all
 CC used in the exemplification of the present invention. The nucleic acid
 CC sequences, protein sequences, and antibodies against them, can be used in
 CC the manufacture of a composition. The composition can be used as a
 CC medicament (or in the manufacture of a medicament) for treating,
 CC preventing or diagnosing infection due to *Neisseria* bacteria. For
 CC example, some of the identified proteins could be components of vaccines
 CC against *Meningococcus B*; against all serotypes; and/or against all
 CC pathogenic *Neisseria*. Identification of sequences from the bacterium
 CC will also facilitate production of biological probes, particularly
 CC organism-specific probes. Attempts to make efficacious *Meningococcus B*
 CC vaccines have failed mainly due to antigen tolerance. Multivalent
 CC antigens have also been tried but none have successfully overcome
 CC antigenic variability. The provision of further, complete sequences may
 CC provide an opportunity to identify secreted or surface exposed proteins
 CC that may be presumed targets for the immune system and which are not
 CC antigenically variable or at least more conserved than other more
 CC variable regions
 XX

SQ Sequence 1437668 BP; 344338A; 353206C; 385074G; 355045T; 0U; 50ther;

Query Match 70.7%; Score 29; DB 3; Length 110000;

Best Local Similarity 100.0%; Pred. No. 1.3;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGAAAAAACAATATTGCGGTATAC 41

Db 79104 TATGAAAAAACAATATTGCGGTATAC 79076

RESULT 4

AAF21544/c

XX ID AAF21544 standard; DNA; 349980 BP.

XX AC AAF21544;

XX DT 13-MAR-2001 (first entry)

XX DE *Neisseria meningitidis* B nucleotide sequence SEQ ID NO.1.

XX KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; immunogenic; vaccine;
 KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
 KW ds.

XX OS *Neisseria meningitidis*.

XX PN WO200066791-A1.

XX PD 09-NOV-2000.

XX PF 08-MAR-2000; 2000WO-US005928.

XX PR 30-APR-1999; 99US-0132068P.

XX PR 08-OCT-1999; 99WO-US023573.
 PR 28-FEB-2000; 2000GB-00004695.

XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.

XX Pizsa M, Hickey B, Peterson J, Tettein H, Venter JC;
 PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V;
 PI Rappoli R, Frazer CM, Grandi G;
 XX

DR WPI; 2000-647603/62.

PT *Neisseria meningitidis* B full length genome sequence and open reading
 PT frames are used to detect, treat and prevent *Neisseria* infections.
 XX

PS Claim 7; Appendix A; 692pp; English.

XX
 CC The present invention describes the full length genome of *Neisseria*
 CC *meningitidis* B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
 CC represent fragments of the NMB genomic sequence, as the sequence was too
 CC long to go in a record on its own it was split into 8 sequences which
 CC overlap each other at the beginning and end of each sequence by 49980 bp
 CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of
 CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
 CC AAF21608, and so on). AAF21545 to AAF21588 encode the *Neisseria* proteins
 CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
 CC primers which are used in the exemplification of the present invention.
 CC The NMB genome and fragments from it have antibacterial activity, and can
 CC be used in vaccines and gene therapy. *Neisseria* nucleic acids, proteins
 CC and/or antibodies which binds to the proteins can be used in compositions
 CC for treating or preventing infection due to *Neisseria* bacteria or as a
 CC diagnostic reagent for detecting the presence of *Neisseria* bacteria or
 CC of antibodies raised to *Neisseria* bacteria. Computers, computer memory,
 CC computer storage medium or computer databases can be used in a search to
 CC identify open reading frames (ORFs) or coding sequences within the NMB
 CC genome. The DNA sequences provide further opportunities to find antigenic
 CC or immunogenic proteins which are more effective in vaccines than the
 CC outer membrane proteins currently used
 XX

SQ Sequence 349980 BP; 83241 A; 85091 C; 95206 G; 86442 T; 0 U; 0 Other;

Query Match 70.7%; Score 29; DB 3; Length 349980;

Best Local Similarity 100.0%; Pred. No. 1.5;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGAAAAAACAATATTGCGGTATAC 41

Db 79104 TATGAAAAAACAATATTGCGGTATAC 79076

RESULT 5

AAZ24472

XX ID AAZ24472 standard; DNA; 687 BP.

XX AC AAZ24472;

XX DT 17-FEB-2000 (first entry)

XX DE *N. meningitidis* siab DNA.

XX KW CMP sialate synthetase; siab; cytidine monophosphate; detection;
 KW CMP-activated N-acetylneuraminic acid; sialyl acceptor; ds.

XX OS *Neisseria meningitidis*.

XX PN DE19913206-A1.

XX PD 07-OCT-1999.

XX PF 24-MAR-1999; 99DE-01013206.

XX PR 26-MAR-1998; 98DE-01013426.

XX PA (FESS/) FESSNER W D.

```
PI Resener W, Knorret M;
XX
XX WPI; 2000-024923/03.
DR P-PSDB; AAY50798.
XX
XX Production of CMP-sialate synthetase protein useful for producing CMP-
PT activated N-acetylneuraminic acid analogs and non-naturally sialylated
XX oligosaccharides and glycoconjugates.
XX
XX Claim 5; Page 19-20; 26pp; German.
XX
XX This invention describes a novel method for the production of a CMP
XX (cytidine monophosphate)-sialate synthetase protein which comprises
XX culturing a prokaryotic host organism transformed with an expression
XX vector containing a structural gene for the protein, where the start
XX codon of the structural gene is 8-12 bases downstream from a ribosome
XX binding site (RBS). The protein is useful for producing CMP-activated N-
XX acetylneuraminic acid analogs by reacting the analogs with cytidine
XX triphosphate in the presence of the protein and for producing non-
XX naturally sialylated oligosaccharides and glycoconjugates, which are of
XX therapeutic interest because of their sialidase resistance. The protein
XX is also useful for detecting sialyl acceptors by reacting the acceptor
XX with a fluorescent sialic acid derivative in the presence of the protein,
XX cytidine triphosphate and a sialyl transferase and detecting the
XX fluorescently modified acceptor by selective chromatography. This
XX sequence encodes a CMP (cytidine monophosphate)-sialate synthetase
XX protein, sIab, isolated from Neisseria meningitidis serogroup B
XX
XX Sequence 687 BP; 218 A; 154 C; 122 G; 193 T; 0 U; 0 Other;
SQ
XX
XX Query Match 68.3%; Score 28; DB 3; Length 687;
XX Best Local Similarity 100.0%; Pred. No. 1.7;
XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 14 ATGGAACAAATATTCGGTTATAC 41
XX |||||
XX 1 ATGGAACAAATATTCGGTTATAC 28
XX
XX
XX RESULT 6
XX AAA53735
XX ID AAA53735 standard; DNA; 41 BP.
XX
XX AC AAA53735;
XX
XX DT 22-DEC-2000 (first entry)
XX
XX DE Primer for amplifying ORF 7a from LPS core biosynthesis locus.
XX
XX KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;
XX acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;
XX Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
XX sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
XX immunity; immunogen; ganglioside; primer; ss.
XX
XX OS Campylobacter jejuni.
XX
XX PN WO200046379-A1.
XX
XX PD 10-AUG-2000.
XX
XX PF 01-FEB-2000; 2000MO-CA000086.
XX
XX PR 01-FEB-1999; 99US-0118213P.
XX
XX PR 31-JAN-2000; 2000US-00495406.
XX
XX (CANA ) NAT RES COUNCIL CANADA.
XX
XX PA Gilbert M, Wakarchuk WW;
XX
XX PI WPI; 2000-524418/47.
XX
XX DR Novel glycosyltransferase polypeptides and polynucleotides useful for
XX
XX PT
```

```
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic
XX reagents and as immunogen for producing antibodies.
XX
XX Example 1; Page 47; 120pp; English.
XX
XX A reaction mixture for the synthesis of a sialylated oligosaccharide is
XX useful for synthesizing sialylated oligosaccharide such as ganglioside,
XX lysoganglioside or their mimics. Glycosyltransferases are useful for
XX chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
XX other oligosaccharides that have biological activity. The enzymes and
XX nucleic acids that encode them are useful for studies of the pathogenesis
XX mechanisms of organisms that synthesize ganglioside mimics, such as C.
XX jejuni and the nucleic acids are used as probes to study expression of
XX genes involved in ganglioside mimetic synthesis. Antibodies raised
XX against the glycosyltransferases are also useful for analyzing the
XX expression patterns of these genes involved in pathogenesis. The nucleic
XX acids are also useful for designing antisense oligonucleotides for
XX inhibiting expression of the campylobacter enzymes that are involved in
XX the biosynthesis of ganglioside mimics that can mask the pathogens from
XX the host's immune system. The oligosaccharides are useful as diagnosing
XX reagents or as therapeutics and as immunogens for producing antibodies.
XX Bacterial glycosyltransferase can be used to catalyse the formation of
XX oligosaccharides that are identical to the corresponding mammalian
XX structures and are easier and less expensive to produce in large
XX quantity, compared to the mammalian glycosyltransferase. The bacterial
XX origin of the enzymes facilitates expression of large quantities of the
XX enzymes using relatively inexpensive prokaryotic expression systems. Two
XX primers (AAA53735, AAA53736) were used to amplify open reading frame 7a
XX from the LPS core biosynthetic locus
XX
XX Sequence 41 BP; 15 A; 3 C; 10 G; 13 T; 0 U; 0 Other;
SQ
XX
XX Query Match 62.0%; Score 25.4; DB 3; Length 41;
XX Best Local Similarity 82.9%; Pred. No. 11;
XX Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 1 CTTAGAGGTCATATGAAAAACAAATATTCGG 35
XX |||||
XX 1 CTTAGAGGTCATATGAAAAACAAATATTCG 35
XX
XX
XX Db 1 CTTAGAGGTCATATGAAAAACAAATATTCG 35
XX
XX
XX RESULT 7
XX ABT13687
XX ID ABT13687 standard; DNA; 41 BP.
XX
XX AC ABT13687;
XX
XX DT 07-FEB-2003 (first entry)
XX
XX DE Campylobacter jejuni bifunctional sialtransferase cslII PCR primer #3.
XX
XX KW PCR; primer; enzyme; gene therapy; acyltransferase; glycosyltransferase;
XX GalNAc transferase; N-Acetylglucosamine transferase; ss;
XX galactosyltransferase; sialyltransferase; sialic acid synthase;
XX cytidine 5'-monophosphate sialic acid synthetase;
XX CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
XX ganglioside mimetics; inflammation; tumour metastasis.
XX
XX OS Campylobacter jejuni.
XX
XX PN WO200274942-A2.
XX
XX PD 26-SEP-2002.
XX
XX PF 22-FEB-2002; 2002MO-CA000229.
XX
XX PR 21-MAR-2001; 2001US-00816028.
XX
XX (CANA ) NAT RES COUNCIL CANADA.
XX
XX PA Gilbert M, Wakarchuk WW;
XX
XX PI WPI; 2003-040554/03.
XX
XX DR
```

XX New glycosyltransferases from *Campylobacter*, useful for synthesizing
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
PT mechanisms of organisms that synthesize ganglioside mimetics.

PS Example, Page 49, 107pp, English.

XX The invention comprises the amino acid and coding sequences of
CC *Campylobacter jejuni* proteins. The C. jejuni proteins of the invention
CC may be either an acyltransferase; glycosyltransferase; GalNAc (N-
CC Acetylglucosamine) transferase; galactosyltransferase;
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
CC sialic acid synthase; acetyltransferase. The C. jejuni DNA and protein
CC sequences of the invention are useful for ganglioside synthesis, studying
CC ganglioside mimetics, and for designing oligonucleotides to inhibit
CC expression of *Campylobacter* enzymes involved in the biosynthesis of
CC ganglioside mimetics that can mask the pathogen's from the host's immune
CC system. The C. jejuni oligosaccharides of the invention may be used as
CC diagnostic reagents (e.g. to locate areas of inflammation or tumor
CC metastasis). The present DNA sequence represents a *Campylobacter jejuni*
CC PCR primer that was used in an example of the invention

XX Sequence 41 BP, 15 A, 3 C, 10 G, 13 T, 0 U, 0 Other;

Query Match 62.0%; Score 25.4; DB 8; Length 41;
Best Local Similarity 82.9%; Pred. No. 11;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATAATTTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAACAATAATTTGCGG 35

RESULT 8
ABK83460
ID ABK83460 standard; cDNA, 165199 BP.

XX ABK83460;

DT 14-AUG-2002 (first entry)

DE Human cDNA differentially expressed in granulocytic cells #31.

XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.

XX Homo sapiens.

XX WO200228999-A2.

PD 11-APR-2002.

PF 03-OCT-2001; 2001WO-US030821.

PR 03-OCT-2000; 2000US-0237189P.

XX (GENB-) GENE LOGIC INC.

PI Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;

XX WPI; 2002-435328/46.

DR Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.
XX Claim 1; SEQ ID NO 31, 114pp; English.
PS

XX The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 165199 BP; 48510 A, 33223 C, 34406 G, 49060 T, 0 U, 0 Other;

Query Match 60.5%; Score 24.8; DB 6; Length 165199;
Best Local Similarity 80.6%; Pred. No. 47;
Matches 29; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 TTAGAGGTCATATGAAAAACAATAATTTGCGGTT 37
DB 113374 TTAGAGGTCATATGAAAAACAATAATTTGCGGTT 113409

RESULT 9
ADO20395

ID ADO20395 standard; cDNA, 4356 BP.

XX ADO20395;

DT 12-AUG-2004 (first entry)

DE Human PRO polynucleotide #643.

XX Human; PRO; gene; ss; immune related disorder;
KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
KW renal disease; demyelinating disease; central nervous system;
KW peripheral nervous system; demyelinating polyneuropathy;
KW Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.

XX Homo sapiens.

XX WO2004043361-A2.

XX 27-MAY-2004.

XX	06-NOV-2003 ; 2003WO-US035268 .
PF	
XX	08-NOV-2002 ; 2002US-0425235P .
PR	
XX	(GETH) GENENTECH INC.
PA	
XX	Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI	Wood WL, Wu TD;
PI	
XX	WPI ; 2004-420067/39 .
DR	P-PSDB ; ADO20396 .
XX	
XX	Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO8638 useful for
PT	treating an immune related disorder such as systemic lupus erythematosus,
PT	rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT	spondyloarthropathy.
XX	
PS	Claim 1 ; SEQ ID NO 1378 ; 1731pp ; English.
XX	
CC	The invention relates to human PRO polypeptides and the polynucleotides
CC	encoding them. The polypeptides and polynucleotides are useful for
CC	treating and diagnosing immune related disorders in mammals. The immune
CC	related disorders include systemic lupus erythematosus, rheumatoid
CC	arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC	sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC	haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC	melittus, immune-mediated renal disease, demyelinating diseases of the
CC	central or peripheral nervous system, demyelinating polyneuropathy,
CC	Gullain-Barre syndrome and chronic inflammatory demyelinating
CC	polyneuropathy. This sequence represents a human PRO polynucleotide of
CC	the invention.
XX	
SQ	Sequence 4356 BP ; 1230 A ; 987 C ; 1027 G ; 1112 T ; 0 U ; 0 Other ;
	Query Match 59.5% ; Score 24.4 ; DB 12 ; Length 4356 ;
	Best Local Similarity 82.4% ; Pred. No. 44 ;
	Matches 28 ; Conservative 0 ; Mismatches 6 ; Indels 0 ; Gaps
OY	2 TTAGGAGGTCATATGAAAAACAAATATTTGCGG 35
Db	650 TTAGGAGGTCATATGAAAAATCAAGACATTGCTG 683
RESULT 10	
ACN90426	
ID	ACN90426 standard ; DNA ; 7183 BP .
XX	
AC	ACN90426 ;
XX	
DT	02-DEC-2004 (first entry)
XX	
DE	Breast cancer related marker, seq id 11576 .
XX	
XX	Cancer ; breast ; tumour ; cytostatic ; marker ; detection ; therapy ; ds .
XX	
OS	Homo sapiens .
XX	
PN	US2003099974-A1 .
XX	
PD	29-MAY-2003 .
XX	
PF	18-JUL-2002 ; 2002US-0019846 .
XX	
PR	18-JUL-2001 ; 2001US-0306220P .
XX	
PA	(MIL-) MILLENNIUM PHARM INC .
XX	
PI	Lillie J, Xu Y, Wang Y, Steinmann K ;
XX	
DR	WPI ; 2003-787014/74 .
XX	
PT	Novel isolated polypeptide associated with breast cancer, useful for

CC to produce mono- or polyclonal antibodies. The polynucleotides,
CC polypeptides, antibodies, vectors, host cells or modulating agents can be
CC used to produce a pharmaceutical composition
XX

SQ Sequence 444 BP; 141 A; 75 C; 99 G; 129 T; 0 U; 0 Other;

Query Match 56.1%; Score 23; DB 3; Length 444;
Best Local Similarity 74.4%; Pred. No. 1.1e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAGAAACAAATATTGCGGTTAT 39
DB 148 CATAGTGTCTTATGCTATATAAAATCATTCGCTTAT 110

RESULT 12

ID ABX65592/C
ID ABX65592 standard; DNA; 1024 BP.

AC ABX65592;

XX 07-MAY-2003 (first entry)

XX Helicobacter pylori selected interacting domain (SID) DNA #191.

XX Protein-protein interaction; ulcer; selected interacting domain; SID;
KM gene; de.

XX Helicobacter pylori.

XX WO200266501-A2.

XX 29-AUG-2002.

XX 28-DEC-2001; 2001WO-EP015428.

XX 02-JAN-2001; 2001US-0259302P.

XX (HYBR-) HYBRIGENICS.

XX (INSP) INST PASTEUR.

XX Legrain P, Rain J, Colland F, De Reuse H, Labigne A;

XX WPI; 2002-674910/72.

XX P-PSDB; ABUS0848.

XX New complexes of protein-protein interactions in Helicobacter pylori,
PT useful for identifying modulating compounds for treating or preventing
PT ulcers in mammals.

XX Claim 7; Page 132; 642pp; English.

XX The invention describes a complex of protein-protein interactions in
CC Helicobacter pylori selected from 421 complexes given in the
CC specification. The complex of protein-protein interactions are useful for
CC screening for agents which modulate the interaction of proteins.

CC Modulating compounds which binds to a targeted bacterial protein may be
CC used for treating or preventing ulcers in a human or animal. This
CC sequence encodes a selected interacting domain (SID), identified via
CC protein-protein interactions

XX Sequence 1024 BP; 344 A; 168 C; 227 G; 285 T; 0 U; 0 Other;

Query Match 56.1%; Score 23; DB 6; Length 1024;
Best Local Similarity 74.4%; Pred. No. 1.2e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAGAAACAAATATTGCGGTTAT 39
DB 150 CATAGTGTCTTATGCTATATAAAATCATTCGCTTAT 112

RESULT 13

ACA34620/C
ID ACA34620 standard; DNA; 1680 BP.

XX ACA34620;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #16277.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KM drug design; gene.

XX Helicobacter pylori.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zykkind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABUS0750.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 22490; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1680 BP; 583 A; 269 C; 339 G; 489 T; 0 U; 0 Other;

Query Match 56.1%; Score 23; DB 8; Length 1680;
 Best Local Similarity 74.4%; Pred. No. 1.3e+02;
 Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CTTAGAGGTCATATGAAAAAACAATATTCGGTTAT 39
 Db 1381 CATAGTGGCTTATGCTATAAAAATCATTCGTTTAT 1343

RESULT 14
 AAX14256/C
 ID AAX14256 standard; DNA; 1780 BP.

AC AAX14256;
 DT 31-MAR-1999 (first entry)
 DE H. pylori GHP0 1488 gene.
 KW GHP0 protein; Helicobacter infection; gastroduodenal disease; gastritis;
 KM peptic ulcer disease; ss.
 XX Helicobacter pylori.
 OS

Key Location/Qualifiers
 FH 51..1730
 FT CDS /*tag= a
 XX

W09843478-A1.

PD 08-OCT-1998.
 PF 01-APR-1998; 98WC-US006371.
 XX

PR 01-APR-1997; 97US-00833457.
 PR 24-JUN-1997; 97US-00881227.
 PR 29-JUL-1997; 97US-00902615.
 XX

PA (INNR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
 PA (HUMA-) HUMAN GENOME SCI INC.

PI Kleantous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;
 DR WPI; 1998-542293/46.
 DR P-PSDB; AAW98537.
 XX

PT New isolated Helicobacter polymnucleotides - used to develop products for
 PT the diagnosis, prevention and treatment of Helicobacter infections and
 PT gastrointestina diseases.
 XX

PS Claim 1; Page 1032-1034; 2054pp; English.

CC This sequence represents a polymnucleotide of the invention. It was
 CC isolated from Helicobacter pylori and encodes a H.pylori GHP0 protein.
 CC The polypeptides can be used for preventing or treating Helicobacter
 CC infections, and gastroduodenal diseases associated with these infections,
 CC including acute, chronic, and atrophic gastritis, and peptic ulcer
 CC diseases, e.g. gastric and duodenal ulcers. They can also be used for the
 CC production of antibodies. The products can also be used for detection and
 CC diagnosis
 XX

SQ Sequence 1780 BP; 623 A; 287 C; 355 G; 515 T; 0 U; 0 Other;

Query Match 56.1%; Score 23; DB 2; Length 1780;
 Best Local Similarity 74.4%; Pred. No. 1.3e+02;
 Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CTTAGAGGTCATATGAAAAAACAATATTCGGTTAT 39
 Db 1431 CATAGTGGCTTATGCTATAAAAATCATTCGTTTAT 1393

RESULT 15

AEA61394
 ID AEA61394 standard; DNA; 4391 BP.

AC AEA61394;

DT 11-AUG-2005 (first entry)

DE Bacillus thuringiensis Cry19Aa coding sequence, SEQ ID 4.

KW Insecticide; crystal protein; Cry19Aa; gene; ds.

OS Bacillus thuringiensis.

Key Location/Qualifiers
 FH 719..2665
 FT CDS /*tag= a
 FT /product= "Cry19Aa"

PN US2005124803-A1.

PD 09-JUN-2005.

PF 30-AUG-2004; 2004US-00929754.

PR 29-AUG-2003; 2003US-0498826P.

PA (OHTS) UNIV OHIO STATE RES FOUND.

PI Dean DH, Abdullah MA;

DR WPI; 2005-417057/42.

DR P-PSDB; AEA61393.

DR EMBL; Y07603.

PT New modified Bacillus thuringiensis insecticidal crystal proteins (i.e.
 PT Cry8Aa and Cry19Aa) with enhanced toxicity, useful for reducing or
 PT eliminating populations of target insects (i.e. mosquitoes) that are
 PT vectors of disease.
 XX

PS Disclosure; SEQ ID NO 4; 63pp; English.

CC The present invention relates to modified insecticidal Bacillus
 CC thuringiensis crystal proteins Cry4Ba and Cry19Aa which have enhanced
 CC toxicity against a variety of insects, e.g. mosquitoes and lepidoptera.
 CC The modified proteins are useful for reducing or eliminating populations
 CC of target insects that are vectors of disease, particularly mosquitoes.
 CC The present sequence is the coding sequence of the wild-type Cry19Aa
 CC protein, which can be modified with: a substitution of amino acids at
 CC positions 355 through 358 with amino acids tyrosine, glutamine, aspartic
 CC acid, and leucine; an insertion of at least one amino acid (e.g.
 CC arginine) after position 358 and a deletion of the amino acids at
 CC positions 414 through 418.
 XX

SQ Sequence 4391 BP; 1557 A; 630 C; 914 G; 1290 T; 0 U; 0 Other;

Query Match 56.1%; Score 23; DB 14; Length 4391;
 Best Local Similarity 74.4%; Pred. No. 1.4e+02;
 Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 3 TTAGAGGTCATATGAAAAAACAATATTCGGTTATAC 41
 Db 1074 TAGAAGATCTTATAGCAAAACGAATTAACGAGTTATAC 1112

Search completed: April 7, 2006, 16:17:25
 Job time : 50.2484 secs

* GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:00:12 ; Search time 321.989 Seconds
(without alignments)
5957.564 Million cell updates/sec

Title: US-09-211-691-3

Perfect score: 41

Sequence: 1 ctctaggagctcatatgtaaa.....acaaatattcggtttatc 41

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	25.8	62.9	638	9	B2245783 CH230-325
c 2	24.6	60.0	583	7	CR750796 DKFZp470M
c 3	24.6	60.0	662	11	CR869763
c 4	24.6	60.0	701	7	CV649651 EST928404
c 5	24.6	60.0	771	7	CV635258 EST914011
c 6	24.6	59.5	500	7	CN361845
c 7	24.4	59.5	521	7	CN361847 170004252
c 8	24.4	59.5	582	3	BP350103
c 9	24.4	59.5	669	3	AL703924 DKFZp686D
c 10	24.4	59.5	790	8	DR001464
c 11	24.4	59.5	821	2	BG756622 602713774
c 12	24.4	59.5	905	2	BQ424531
c 13	24.4	59.5	1113	3	BM557407 AGENCCOURT
c 14	24.4	59.5	4191	10	AY399307
c 15	24.4	59.5	4503	10	AY399306
c 16	24.2	59.0	672	9	BH298686 CH230-84H
c 17	24.2	59.0	1190	3	BI687776
c 18	24.2	58.5	439	2	BI065289
c 19	24.2	58.5	531	1	AJ741063
c 20	24.2	58.5	619	5	BU296809
c 21	24.2	58.5	695	9	BH951067
c 22	24.2	58.5	733	5	BU200026

c 23	24	58.5	773	9	BH487940
c 24	23.6	57.6	234	8	DR176547
c 25	23.6	57.6	699	2	BF681440
c 26	23.6	57.6	789	9	BZ265368
c 27	23.6	57.6	890	10	AG424476
c 28	23.4	57.1	261	1	AV164171
c 29	23.4	57.1	455	6	CF533995
c 30	23.4	57.1	546	5	BU301596
c 31	23.4	57.1	591	9	AZ382739
c 32	23.4	57.1	654	6	CB521772
c 33	23.4	57.1	658	1	AM064320
c 34	23.4	57.1	731	5	BU056866
c 35	23.4	57.1	806	6	CD350955
c 36	23.4	57.1	4474	10	AY399308
c 37	23.2	56.6	530	9	BZ443595
c 38	23.2	56.6	658	7	CJ392627
c 39	23.2	56.6	710	10	EX171933
c 40	23.2	56.6	910	10	AG893009
c 41	23.2	56.6	927	7	CJ403240
c 42	23	56.1	265	1	AJ752805
c 43	23	56.1	265	1	AJ752815
c 44	23	56.1	503	7	CR455290
c 45	23	56.1	519	11	DR19H65

ALIGNMENTS

RESULT 1
LOCUS B2245783/c
DEFINITION CH230-325M4.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
ACCESSION B2245783
VERSION B2245783.1 GI:23960407
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
1 (bases 1 to 638)
Zhao,S., Shetty,J., Shatsman,S., Tsagaye,G., Geer,K., Rigg,B., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other GSSs: CH230-325M4.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.choi.org/bacpac/rat230.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.choi.org).
Clones may be purchased from BACPAC Resources
(http://www.choi.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tdb/bac_end/rat/bac_end_intro.html
Plate: 325 row: M column: 4
Seq primer: SPE
Clase: BAC ends.
Location/Qualifiers
1. 638
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SNHsd/MC9"
/db_xref="taxon:10116"
/clone="CH230-325M4"
/sex="Female"
/cell_type="Brain"

ORIGIN

/clone.lib="CHORI-230 Segment 2"
/note=Vector: PITABACT.3; Site_1: Mbol; Site_2: Mbol;
CHORI-230 Rat (BN/SENhd/MCW) BAC library produced by
Pieter de Jong"

Query Match 62.9%; Score 25.8; DB 9; Length 638;
Best Local Similarity 93.1%; Pred. No. 1.4e+02;
Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 9 GTCATATGGAACAAATATTGCGTT 37

Db 287 GTCATATGGAACAAATATTGCGTT 259

RESULT 2

LOCUS

CR750796 583 bp mRNA linear EST 30-AUG-2004
DKFZP470M2222.r1 470 (synonym: pliv1) Pongo pygmaeus CDNA clone
DKFZP470M2222.5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 583)
Pousotka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Mewes, H.W., Well, B., Amid, C., Osanger, A., Fodor, G., Han, M. and
Wiemann, S.
Pongo pygmaeus mRNA (Pousotka, A., Albert, R., Moosmayer, P., et al.)
Unpublished (2004)
Contact: MIPS

AUTHORS

TITLE
JOURNAL
COMMENT

MIPS

Ingolsteader landstr.1, D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email: s.wiemann@dkfz-heidelberg.de; any. Please contact RPD for
ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZP470M2222
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES

source

1..583
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZP470M2222"
/issue_type="liver"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="470 (synonym: pliv1)"
/note=Vector: pSport1_sfi; Site_1: SfiI; Site_2: SfiI"

ORIGIN

Query Match 60.0%; Score 24.6; DB 7; Length 583;
Best Local Similarity 76.9%; Pred. No. 3.7e+02;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 2 TTAGGAGTCATATGGAACAAATATTGCGTTATA 40

Db 545 TTATGAGGAGTCATGGAACAAATATTGCGTTATA 583

RESULT 3

LOCUS

CR869763 662 bp DNA linear GSS 19-NOV-2004

DEFINITION

ACCESSION

VERSION

KEYWORDS

CR869763 Sus scrofa BES, genomic survey sequence.
CR869763 CR869763.1 GI:55868009
GSS; Bac-end sequence BES; Genome Survey Sequence.

SOURCE

ORGANISM

Sus scrofa (pig)
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suidae;
Sus.

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

1 (bases 1 to 662)
Rogel-Gallard, C., Bourgeaux, N., Billault, A., Vaiman, M. and
Chardon, P.
Construction of a swine BAC library: application to the
characterization and mapping of porcine type C endoviral elements
Cytochrome. Cell Genet. 85 (3-4), 205-211 (1999)
10449899
2 (bases 1 to 662)
Chardon, P., Iannucci, N., Rogel, A., Dossat, C., Demars, J.,
Rogel-Gallard, C., Roy, A., Schibler, L. and Milan, D.
A physical map of the swine genome
Unpublished
3 (bases 1 to 662)
Genoscope.
Direct Submission
Submitted (18-NOV-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Location/Qualifiers
1..662
/organism="Sus scrofa"
/mol_type="genomic DNA"
/strain="Large White"
/db_xref="taxon:9823"
/clone="b10733E03"
/sex="male"
/cell_type="fibroblast"
/clone_lib="SBAB"
/note="Genoscope sequence ID : IH0AA72CG01FM1"

/db xref="taxon:5855"
/clone="PVMHPS4"
/note="Vector: Lambda Triplex2; Site 1: Sfi 1A; Site 2: Sfi 1B; Plasmodium vivax field isolate cDNA library made in Lambda Triplex2. Inserts cloned unidirectionally in the Sfi 1A and Sfi 1B sites. Mass excision of library produced inserts in Triplex2 plasmid. Inserts sequenced from either 5' or 3' end using Triplex2 sequencing primer or polydr 24 bp primer respectively."

ORIGIN

Query Match 60.0%; Score 24.6; DB 7; Length 701;
Best Local Similarity 87.1%; Pred. No. 3.8e+02;
Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 AGAGGTCATATGAAAAAATAATTTGCG 34
|||||
DB 379 AGGAGTTCATCTGAAAAAATAATTTGCG 409
|||||

RESULT 5
CN3635258 771 bp mRNA linear EST 01-JAN-2005
DEFINITION EST914011 Field isolate cDNA library Plasmodium vivax cDNA clone
ACCESSION PVMBK89 5' end, mRNA sequence.
VERSION CV635258
KEYWORDS CV635258.1 GI:56942076
SOURCE EST.
ORGANISM Plasmodium vivax (malaria parasite P. vivax)
Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 771)
Carlton, J.M., and Cui, L.
A survey of genes in Plasmodium vivax by EST sequencing
Unpublished (2004)
Contact: Jane Carlton
Parasite Genomics Group
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-530-9319
Fax: 301-838-0208
Email: carlton@tigr.org
Seq primer: TI.

FEATURES
source
Location/Qualifiers
1..771
/organism="Plasmodium vivax"
/mol_type="mRNA"
/strain="Field isolate"
/db_xref="taxon:5855"
/clone="PVMK89"
/clone_lib="Field isolate cDNA library"
/note="Vector: Lambda Triplex2; Site 1: Sfi 1A; Site 2: Sfi 1B; Plasmodium vivax field isolate cDNA library made in Lambda Triplex2. Inserts cloned unidirectionally in the Sfi 1A and Sfi 1B sites. Mass excision of library produced inserts in Triplex2 plasmid. Inserts sequenced from either 5' or 3' end using Triplex2 sequencing primer or polydr 24 bp primer respectively."

ORIGIN

Query Match 60.0%; Score 24.6; DB 7; Length 771;
Best Local Similarity 87.1%; Pred. No. 3.8e+02;
Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 AGAGGTCATATGAAAAAATAATTTGCG 34
|||||
DB 185 AGGAGTTCATCTGAAAAAATAATTTGCG 215
|||||

RESULT 6
CN361845 500 bp mRNA linear EST 16-MAY-2004
LOCUS CN361845

DEFINITION 17000424189354 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN361845
VERSION CN361845.1 GI:47361779
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 500)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 500 Std Error: 0.00.
Location/Qualifiers
1..500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/ligase_type="embryonic stem cells, embryoid bodies derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="Oligo dT primed, full-length enriched cDNA library from embryoid body outgrowths derived from h9s cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."

ORIGIN

Query Match 59.5%; Score 24.4; DB 7; Length 500;
Best Local Similarity 82.4%; Pred. No. 4.3e+02;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TTAGAGGTCATATGAAAAAATAATTTGCG 35
|||||
DB 326 TTAGAGGTCATATGAAAAAATAATTTGCG 359
|||||

RESULT 7
CN361847 521 bp mRNA linear EST 16-MAY-2004
LOCUS CN361847
DEFINITION 17000425249082 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN361847
VERSION CN361847.1 GI:47361781
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 521)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658

Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 521 Std Error: 0.00.
Location/Qualifiers

FEATURES

source

1..521
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="embryonic stem cells, cell lines H1, H7, and H9"
/clone_id="GRN ES"
/note="Toigo dr primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN

Query Match 59.5%; Score 24.4; DB 7; Length 521;
Best Local Similarity 82.4%; Pred. No. 4.4e+02;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy

2 TTAGAGGTCATATGGAACAAATATTGCGG 35
|||||
289 TTAGAGGTCATATGGAACAAATCAAGACATTGCTG 322

RESULT 8 582 bp mRNA linear EST 17-SEP-2004
BP350103 Sugano cDNA library, brain Homo sapiens cDNA clone

LOCUS

BP350103 SZR07565, mRNA sequence.

ACCESSION

BP350103

VERSION

BP350103.1 GI:52280089

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 582)

Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,

Mizushima-Sugano, U., Nakai, K. and Sugano, S.

Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

Genome Res. 14 (9), 1711-1718 (2004)

15342556

Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp.

Location/Qualifiers

1..582

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="SZR07565"

/issue_type="brain"

/clone_id="Sugano cDNA library, brain"

ORIGIN

Query Match 59.5%; Score 24.4; DB 3; Length 582;
Best Local Similarity 82.4%; Pred. No. 4.4e+02;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy

2 TTAGAGGTCATATGGAACAAATATTGCGG 35
|||||
549 TTAGAGGTCATATGGAACAAATCAAGACATTGCTG 582

LOCUS

AL703924 669 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686D1628_r1 666 (synonym: hlc3) Homo sapiens cDNA clone

DKFZp686D1628 5', mRNA sequence.

AL703924
AL703924.1 GI:19687279

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

1 (bases 1 to 669)

Ostenwelder, B., Obermaier, B., Mewes, W., Mewes, H.W., Weil, B. and

Wiemann, S.

EST (Ostenwelder, B., Obermaier, B., Mewes, H.W., Weil, B. and

Wiemann, S.)

Unpublished (2001)

Contact: MIPS

MIPS

Ingo Straeter Landstr. 1, D-85764 Neuberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by Medigenomix (Martinried/Germany) within the cDNA

sequencing consortium of the German Genome Project. No 81 sequence

available.

This clone (DKFZp686D1628) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Location/Qualifiers

1..669

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZp686D1628"

/dev stage="adult"

/lab_host="DH10B"

/clone_id="666 (synonym: hlc3)"

/note="Vector: pTriplEx2, Site_1: SfiI, Site_2: SfiI, B;

cDNA-collection"

ORIGIN

Query Match 59.5%; Score 24.4; DB 1; Length 669;
Best Local Similarity 82.4%; Pred. No. 4.4e+02;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy

2 TTAGAGGTCATATGGAACAAATATTGCGG 35
|||||
546 TTAGAGGTCATATGGAACAAATCAAGACATTGCTG 579

LOCUS

DR001464 790 bp mRNA linear EST 17-MAY-2005
DEFINITION TC114916 Human fetal brain, large insert, pCMV expression library

Homo sapiens cDNA clone TC114916 5' similar to Homo sapiens

Vpr-binding protein (VprBP), mRNA sequence.

DR001464

DR001464.1 GI:66261337

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

1 (bases 1 to 790)

Birkett, C., Cho, J., Gau, Y., Hamer, R., Kelly, S., Kovacs, K., Liu, L.,

Liu, X., Porter, J., Sachs, A., Shu, Y., Sun, Z., Wong, J., Wu, M.,

Zhang, X., Jay, G. and He, W.

High-throughput cloning of full-length human cDNAs directly from

cDNA libraries optimized for large and rare transcripts

Unpublished (2005)

Contact: Kovacs, KF

High Throughput cDNA Cloning

OrGene Technologies, Inc. (www.origene.com)

6 Taft Court, Suite 100, Rockville, MD 20850, USA
 Tel: 301 340 3188
 Fax: 301 340 8606
 Email: CDNA@origene.com
 This EST submission is part of an on-going human full-length
 cloning project at Origene Technologies, Inc.
 Please contact Origene for access.
 Origene Technologies, Inc.
 6 Taft Ct. Suite 100
 Rockville, MD 20850
 Tel: (301) 340-3188
<http://www.origene.com>
 Seq primer: PCW6 5prime forward vector primer, Origene
 Technologies Inc.

FEATURES

source
 1..790
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="TC114916"
 /tissue_type="Fetal Brain"
 /clone_lib="Human fetal brain, large insert, PCW
 expression library"
 /note="Organ: Fetal Brain; Vector: PCW6-XL4; Site 1:
 EcoRI; Site 2: XhoI/Sall compatible end ligatio; Oligo-dT
 *, primed reverse transcription optimized for large and GC
 rich mRNA transcripts, cDNA size selection, optimized
 ligation for large inserts into mammalian expression
 vector, random clones selected for end sequence
 verification of full-length genes"

ORIGIN

Query Match 59.5%; Score 24.4; DB 8; Length 790;
 Best Local Similarity 82.4%; Pred. No. 4.4e+02;
 Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 2 TTAGAGGTCATATGGAACAAATATTCGCG 35
 585 TTAGAGGTCATATGGAACAAATCAAGCATTCGCG 618

RESULT 11
 BG756622 821 bp mRNA linear EST 15-MAY-2001
 LOCUS 602713774P1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853730 5',
 DEFINITION mRNA sequence.
 ACCESSION BG756622
 VERSION BG756622 GI:14067275
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 NIH-MGC <http://mgc.nci.nih.gov/>
 1 (bases 1 to 821)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1CM1700 row: d column: 19
 High quality sequence stop: 790.
 Location/Qualifiers
 1..821
 /organism="Homo sapiens"
 /mol_type="mRNA"

/db_xref="taxon:9606"
 /clone="IMAGE:4853730"
 /tissue_type="Primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_48"
 /note="Organ: B-cells; Vector: pOTR7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 59.5%; Score 24.4; DB 2; Length 821;
 Best Local Similarity 82.4%; Pred. No. 4.5e+02;
 Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TTAGAGGTCATATGGAACAAATATTCGCG 35
 650 TTAGAGGTCATATGGAACAAATCAAGCATTCGCG 683

RESULT 12
 BQ424531 905 bp mRNA linear EST 23-MAY-2002
 LOCUS AGENCOURT_7907296 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6154690
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ424531
 VERSION BQ424531 GI:21119846
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1AM13496 row: c column: 11
 High quality sequence stop: 683.
 Location/Qualifiers
 1..905
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6154690"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_67"
 /note="Organ: eye; Vector: PCW6-SPORT6; Site 1: NotI;
 Site 2: Sall; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

FEATURES

source
 1..905
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6154690"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_67"
 /note="Organ: eye; Vector: PCW6-SPORT6; Site 1: NotI;
 Site 2: Sall; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

ORIGIN
 Query Match 59.5%; Score 24.4; DB 5; Length 905;
 Best Local Similarity 82.4%; Pred. No. 4.5e+02;
 Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 2 TTAGAGGTCATATGGAACAAATATTCGCG 35

Db 217 TTAGAGGTCATATGGAATAATCAAGACATTGCTG 250

RESULT 13
LOCUS BM557407 1113 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT 6561799 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5547856
5', mRNA sequence.
ACCESSION BM557407
VERSION BM557407.1 GI:18799373
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1113)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgrabbs-remail.nih.gov
Tissue Procurement: ATCC/DCMP/DMP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINT at:
http://image.jnl.gov
Plate: L1AM1225 row: n column: 17
High quality sequence step: 579.
Location/Qualifiers
1. .1113
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5547856"
/issue_type="melanotic melanoma"
/lab_host="PH10B (phage-resistant)"
/clone_id="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN
.\$
Query Match 59.5%; Score 24.4; DB 3; Length 1113;
Best Local Similarity 82.4%; Pred. No. 4.5e+02;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TTAGAGGTCATATGGAATAATCAAGACATTGCTG 35
|||||
576 TTAGAGGTCATATGGAATAATCAAGACATTGCTG 609
|||||

RESULT 14
LOCUS AY399307 4191 bp DNA linear GSS 12-DEC-2003
DEFINITION Pan troglodytes HOMO180 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY399307
VERSION AY399307.1 GI:39755295
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.
1 (bases 1 to 4191)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Slnsky,J.J.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Slnsky,J.J.,

TITLE Adams,M.D. and Cargill,M.
Interfing nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 4191)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Slnsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. .4191
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1. .>4191
/locus_tag="HOMO180"

ORIGIN
gene

Query Match 59.5%; Score 24.4; DB 10; Length 4191;
Best Local Similarity 82.4%; Pred. No. 3e+02;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TTAGAGGTCATATGGAATAATCAAGACATTGCTG 35
|||||
451 TTAGAGGTCATATGGAATAATCAAGACATTGCTG 484
|||||

Db 451 TTAGAGGTCATATGGAATAATCAAGACATTGCTG 484
|||||

RESULT 15
LOCUS AY399306 4503 bp DNA linear GSS 12-DEC-2003
DEFINITION Homo sapiens HOMO180 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY399306
VERSION AY399306.1 GI:39755295
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 4503)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Slnsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Interfing nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 4503)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Slnsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. .4503
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1. .>4503
/locus_tag="HOMO180"

ORIGIN
gene

Query Match 59.5%; Score 24.4; DB 10; Length 4503;
 Best Local Similarity 82.4%; Pred. No. 5e+02; 6; Indels 0; Gaps 0;
 Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TTAGAGGTGATATGAAAAACAATATTGCGG 35
 DB 451 TTAGAGGTGCTATGAAAAATCAAGACATTGCTG 484

Search completed: April 7, 2006, 18:17:11
 Job time : 325.989 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:17:47 ; Search time 14.6685 Seconds
(without alignments)
4968.475 Million cell updates/sec

Title: US-09-211-691-3

Perfect score: 41

Sequence: 1 cttagagagtcatactgaa.....acaaatattgcgttac 41

Scoring table: IDENTITY_NUC

Gapd 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/1/COMB.seq:*
2: /cgn2_6/prodata/1/ina/5/COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A/COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B/COMB.seq:*
5: /cgn2_6/prodata/1/ina/H/COMB.seq:*
6: /cgn2_6/prodata/1/ina/PCTUS/COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP/COMB.seq:*
8: /cgn2_6/prodata/1/ina/RE/COMB.seq:*
9: /cgn2_6/prodata/1/ina/backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25.4	62.0	41	US-09-495-406-32	Sequence 32, Appl
2	25.4	62.0	41	US-09-816-028A-46	Sequence 46, Appl
3	25.4	62.0	41	US-10-303-162-46	Sequence 46, Appl
4	25.4	62.0	41	US-10-303-134-46	Sequence 46, Appl
5	25.4	62.0	41	US-10-303-118-46	Sequence 46, Appl
6	25.4	62.0	41	US-10-303-128-46	Sequence 46, Appl
7	24.6	60.0	41	US-09-949-016-12707	Sequence 12707, A
8	24.6	60.0	41	US-09-949-016-17026	Sequence 17026, A
9	23	56.1	444	US-10-012-819-93	Sequence 93, Appl
10	22.4	54.6	451924	US-09-949-016-12896	Sequence 12896, A
11	22.4	54.6	451924	US-09-949-016-17305	Sequence 17305, A
12	22	53.7	601	US-09-949-016-196368	Sequence 196368, A
13	22	53.7	1401	US-09-328-352-4116	Sequence 4116, Ap
14	22	53.7	1422	US-09-107-532A-2287	Sequence 2287, Ap
15	21.8	53.2	601	US-09-949-016-47720	Sequence 47720, A
16	21.8	53.2	6373	US-08-462-728-1	Sequence 1, Appl
17	21.8	53.2	6373	US-08-461-917-1	Sequence 1, Appl
18	21.8	53.2	6373	US-08-464-436-1	Sequence 1, Appl
19	21.8	53.2	6373	US-08-464-436-1	Sequence 1, Appl
20	21.8	53.2	6375	US-08-168-917-5	Sequence 5, Appl
21	21.8	53.2	6375	US-08-460-510-5	Sequence 5, Appl
22	21.8	53.2	6375	US-08-460-510-5	Sequence 5, Appl
23	21.8	53.2	6375	PCT-US92-00730-5	Sequence 5, Appl
24	21.8	53.2	6375	PCT-US92-00862-5	Sequence 5, Appl

C	25	21.8	53.2	58829	3	US-09-949-016-13146	Sequence 13146, A
	26	21.8	53.2	331814	3	US-09-949-016-12008	Sequence 12008, A
	27	21.8	53.2	331814	3	US-09-949-016-17056	Sequence 17056, A
	28	21.6	52.7	65848	3	US-09-949-016-13285	Sequence 13285, A
	29	21.4	52.2	2838	3	US-09-543-681A-28	Sequence 28, Appl
	30	21.4	52.2	53558	3	US-09-949-016-16616	Sequence 16616, A
	31	21.4	52.2	63187	3	US-09-949-016-12682	Sequence 12682, A
	32	21.4	52.2	63187	3	US-09-949-016-16288	Sequence 16288, A
	33	21.4	52.2	197331	3	US-09-949-016-13675	Sequence 12675, A
	34	21.4	52.2	197332	3	US-09-949-016-17170	Sequence 17170, A
	35	21	51.2	789	6	PCT-US96-05350A-1275	Sequence 1275, Ap
	36	21	51.2	837	3	US-09-134-000C-2984	Sequence 2984, Ap
	37	21	51.2	1218	3	US-09-543-681A-2852	Sequence 2852, Ap
	38	21	51.2	2585	3	US-09-573-080A-8	Sequence 8, Appl
	39	21	51.2	26076	3	US-09-949-016-12004	Sequence 12004, A
	40	21	51.2	26076	3	US-09-949-016-13041	Sequence 13041, A
	C 41	21	51.2	52496	3	US-09-949-016-16118	Sequence 16118, A
	C 42	21	51.2	52496	3	US-09-949-016-16119	Sequence 16119, A
	C 43	21	51.2	640681	3	US-09-790-988-1	Sequence 1, Appl
	C 44	21	51.2	786431	3	US-09-751-389-3	Sequence 3, Appl
	45	21	51.2	1830121	3	US-09-557-884-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-495-406-32
; Sequence 32, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Ganglioside and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
US-09-495-406-32
Query Match 62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
C
1 CTTAGAGGTCATATGAAAAAACAATATTCGGG 35
DB 1 CTTAGAGGTCATATGAAAAAAGTTATTCGCTG 35
RESULT 2
US-09-816-028A-46
; Sequence 46, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Ganglioside and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/816,028A

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
; OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-09-816-028A-46

Query Match      62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
Db      1 CTTAGAGGTCATATGAAAAAGTTATTTGCTG 35

RESULT 3
US-10-303-162-46
; Sequence 46, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
; OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-10-303-162-46

Query Match      62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
Db      1 CTTAGAGGTCATATGAAAAAGTTATTTGCTG 35

RESULT 4
US-10-303-134-46
; Sequence 46, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
```

```

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
; OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-10-303-134-46

Query Match      62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
Db      1 CTTAGAGGTCATATGAAAAAGTTATTTGCTG 35

RESULT 5
US-10-303-118-46
; Sequence 46, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
; OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-10-303-118-46

Query Match      62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
Db      1 CTTAGAGGTCATATGAAAAAGTTATTTGCTG 35

RESULT 6
US-10-303-128-46
; Sequence 46, Application US/10303128
; Patent No. 6911337
```



```

; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Makarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Cl-131 5'
; OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-10-303-128-46

```

```

Query Match          62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 1 CTTAGAGGTCATATGAAAAACAAATATTTGCGG 35
Db 1 CTTAGAGGTCATATGAAAAAGTATATTTGCTG 35

RESULT 7
US-09-949-016-12707
; Sequence 12707, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12707
; LENGTH: 190078
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(190078)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12707

```

```

Query Match          60.0%; Score 24.6; DB 3; Length 190078;
Best Local Similarity 76.9%; Pred. No. 12;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy 2 TTAGAGGTCATATGAAAAACAAATATTTGCGGTTATA 40
Db 45703 TTAGAGGTCATATGATATATTCACAGTTGCGTTATA 45741

```

```

RESULT 8
US-09-949-016-17026
; Sequence 17026, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17026
; LENGTH: 190078
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(190078)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17026

```

```

Query Match          60.0%; Score 24.6; DB 3; Length 190078;
Best Local Similarity 76.9%; Pred. No. 12;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy 2 TTAGAGGTCATATGAAAAACAAATATTTGCGGTTATA 40
Db 45703 TTAGAGGTCATATGATATATTCACAGTTGCGGTTATA 45741

```

```

RESULT 9
US-10-012-819-93/C
; Sequence 93, Application US/10012819
; Patent No. 6916615
; GENERAL INFORMATION:
; APPLICANT: Legrain, Pierre
; APPLICANT: Selig, Luc
; TITLE OF INVENTION: Collection of Prokaryotic DNA for Two-Hybrid System, Helicobacter
; FILE REFERENCE: B5053
; CURRENT APPLICATION NUMBER: US/10/012,819
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: EP 99401066.8
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
; LENGTH: 444
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(444)
US-10-012-819-93

```

```

Query Match          56.1%; Score 23; DB 3; Length 444;
Best Local Similarity 74.4%; Pred. No. 25;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

Qy 1 CTTAGAGGTCATATGAAAAACAAATATTTGCGGTTAT 39
Db 148 CATAAGTGCTTATGATATTAATAAATCATTCGCTTAT 110

```

```
RESULT 10
US-09-949-016-12896
; Sequence 12896, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12896
; LENGTH: 451924
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12896
```

```
Query Match          54.6%; Score 22.4; DB 3; Length 451924;
Best Local Similarity 81.2%; Pred. No. 82;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 9 GTCATATGGAAAAACAATATTCGGCTTATA 40
DB 303773 GTCATATGGAAAAACAATATTCGGCTTATA 303804
```

```
RESULT 11
US-09-949-016-17305
; Sequence 17305, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17305
; LENGTH: 451925
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17305
```

```
Query Match          54.6%; Score 22.4; DB 3; Length 451925;
Best Local Similarity 81.2%; Pred. No. 82;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 9 GTCATATGGAAAAACAATATTCGGCTTATA 40
DB 303773 GTCATATGGAAAAACAATATTCGGCTTATA 303804
```

```
RESULT 12
US-09-949-016-196368
; Sequence 196368, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 196368
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-196368
```

```
Query Match          53.7%; Score 22; DB 3; Length 601;
Best Local Similarity 73.7%; Pred. No. 60;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 2 TTAGAGGTCATATGAAAAACAATATTCGGCTTAT 39
DB 517 TTAGGAATCATGTTGGAAAAACAATATTTGCAATGT 554
```

```
RESULT 13
US-09-328-352-4116
; Sequence 4116, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: GARY L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4116
; LENGTH: 1401
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4116
```

```
Query Match          53.7%; Score 22; DB 3; Length 1401;
Best Local Similarity 73.7%; Pred. No. 66;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 2 TTAGAGGTCATATGAAAAACAATATTCGGCTTAT 39
DB 667 TTAGTTCGTCATATGAAAAACAATGCTGATGT 704
```

```
RESULT 14
US-09-107-532A-2287
; Sequence 2287, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSES: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: CD/ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 2287:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1422 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (8) LOCATION 1..1422
 SEQUENCE DESCRIPTION: SEQ ID NO: 2287:
 US-09-107-532A-2287

Query Match 53.7%; Score 22; DB 3; Length 1422;
 Best Local Similarity 83.3%; Pred. No. 66;
 Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 7 AGGTCATATGAAAAACAATAATTGCGGT 36
 Db 378 AGAGCATCTGAAAAACAATAATTGCGTT 407

RESULT 15
 US-09-949-016-47720
 ; Sequence 47720, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 47720
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-47720

Query Match 53.2%; Score 21.8; DB 3; Length 601;
 Best Local Similarity 70.7%; Pred. No. 71;
 Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 CTTAGAGGTCATATGAAAAACAATAATTGCGTTATAC 41
 Db 28 CTTAGAAAGTCATATGAAAAACAATAATTGCGTTATAC 68

Search completed: April 7, 2006, 18:22:43
 Job time : 18.6685 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using SW model

Run on: April 7, 2006, 18:02:11 ; Search time 67.1969 Seconds
(without alignments)
5045.535 Million cell updates/sec

Title: US-09-211-691-3

Perfect score: 41

Sequence: 1 cttagagagtcatactgaa.....acaaatctgcgtrtac 41

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA_Main:*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
8: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	41	3	US-09-211-691-3
2	41	100.0	41	6	US-10-317-773-3
3	41	100.0	41	6	US-10-317-773-3
4	29	70.7	65632	9	US-10-915-740A-49
5	29	70.7	2242716	9	US-10-915-740A-1068
6	25.4	62.0	41	3	US-09-816-028A-46
7	25.4	62.0	41	6	US-10-303-161-46
8	25.4	62.0	41	6	US-10-303-118-46
9	25.4	62.0	41	6	US-10-303-128-46
10	25.4	62.0	41	6	US-10-303-134-46
11	25.4	62.0	41	6	US-10-303-162-46
12	25.4	62.0	41	8	US-10-735-419-46
13	25.4	62.0	41	8	US-10-820-536-46
14	25.4	62.0	41	8	US-10-845-408-46
15	25.4	62.0	41	8	US-10-845-412-46
16	25.4	62.0	41	8	US-10-846-219-46
17	25.4	62.0	41	8	US-10-821-604-46
18	25.4	62.0	41	8	US-10-847-983-46
19	25.4	62.0	41	8	US-10-821-573-46
20	25.4	62.0	41	8	US-10-850-807-46
21	25.4	62.0	41	8	US-10-850-125-46
22	25.4	62.0	41	8	US-10-830-825-46
23	25.4	62.0	41	9	US-10-962-334-46

24	25.4	62.0	41	9	US-10-830-997-46	Sequence 46, Appl
25	25.4	62.0	41	9	US-10-962-235-46	Sequence 46, Appl
26	25.4	62.0	41	9	US-10-961-882-46	Sequence 46, Appl
27	24.4	59.5	7183	5	US-10-198-846-11576	Sequence 11576, A
28	23.2	56.6	529	4	US-09-925-065A-733394	Sequence 733394, A
29	23	56.1	444	5	US-10-012-819-93	Sequence 93, Appl
30	23	56.1	1680	7	US-10-282-122A-22490	Sequence 22490, A
31	23	56.1	1780	3	US-09-895-912A-103	Sequence 103, App
32	23	56.1	4391	9	US-10-929-754-4	Sequence 4, Appl
33	23	56.1	3673778	6	US-10-312-841-2	Sequence 2, Appl
34	22.8	55.6	1095	4	US-09-925-065A-723641	Sequence 723641, A
35	22.8	55.6	1095	4	US-09-925-065A-723642	Sequence 723642, A
36	22.8	55.6	1095	4	US-09-925-065A-723643	Sequence 723643, A
37	22.6	55.1	1538	7	US-10-398-221-3367	Sequence 3367, Ap
38	22.6	55.1	14012	9	US-09-819-994-3	Sequence 3, Appl
39	22.6	55.1	14012	9	US-10-473-340-3	Sequence 3, Appl
40	22.4	54.6	528	4	US-09-925-065A-722562	Sequence 272562, A
41	22.4	54.6	528	4	US-09-925-065A-722563	Sequence 272563, A
42	22.4	54.6	528	4	US-09-925-065A-722564	Sequence 272564, A
43	22.4	54.6	575	4	US-09-925-065A-879344	Sequence 879344, A
44	22.4	54.6	575	4	US-09-925-065A-907561	Sequence 907561, A
45	22.4	54.6	578	4	US-09-925-065A-876021	Sequence 876021, A

ALIGNMENTS

RESULT 1
US-09-211-691-3
Sequence 3, Application US/09211691
Patent No. US20020034805A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Fusion Proteins for Use in Enzymatic Synthesis of
FILE REFERENCE: 019957-012910US
CURRENT APPLICATION NUMBER: US/09/211,691
CURRENT FILING DATE: 1998-12-14
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:SYNTH-P1 5'
US-09-211-691-3
Query Match
Best Local Similarity 100.0%; Score 41; DB 3; Length 41;
Matches 41; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 1 CTTAGAGGTCATATGAGAAACAAATATTCGGTTATAC 41
DB 1 CTTAGAGGTCATATGAGAAACAAATATTCGGTTATAC 41
RESULT 2
US-10-317-773-3
Sequence 3, Application US/10317773
Patent No. US20030180928A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Fusion Proteins for Use in Enzymatic Synthesis of

*No Dhl
Cloning strategy to a
different fusion
protein*

TITLE OF INVENTION: GalNAc Transferase
FILE REFERENCE: 019633-000812US
CURRENT APPLICATION NUMBER: US/10/317,773
CURRENT FILING DATE: 2002-12-11
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
PRIOR APPLICATION NUMBER: US 09/211,691
PRIOR FILING DATE: 1998-12-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence:SYNTH-F1 5'
OTHER INFORMATION: primer
US-10-317-773-3

Query Match 100.0%; Score 41; DB 6; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAGAAACAAATATTGCGGTATAC 41
DB 1 CTTAGAGGTCATATGAGAAACAAATATTGCGGTATAC 41

RESULT 3
US-10-317-428-3
Sequence 3, Application US/10317428
Publication No. US20030186414A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Young, N. Martin
APPLICANT: Makarchuk, Warren W.
APPLICANT: National Research Council of Canada
TITLE OF INVENTION: Nucleic Acid That Encodes a Fusion Protein
FILE REFERENCE: 019633-000811US
CURRENT APPLICATION NUMBER: US/10/317,428
CURRENT FILING DATE: 2002-12-11
PRIOR APPLICATION NUMBER: US 60/069,443
PRIOR FILING DATE: 1997-12-15
PRIOR APPLICATION NUMBER: US 09/211,691
PRIOR FILING DATE: 1998-12-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence:SYNTH-F1 5'
OTHER INFORMATION: primer
US-10-317-428-3

same as above

Query Match 100.0%; Score 41; DB 6; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAGAAACAAATATTGCGGTATAC 41
DB 1 CTTAGAGGTCATATGAGAAACAAATATTGCGGTATAC 41

RESULT 4
US-10-915-740A-49
Sequence 49, Application US/10915740A
Publication No. US20050191316A1
GENERAL INFORMATION:
APPLICANT: Frazer, Claire M.
APPLICANT: Hickey, Erin
APPLICANT: Peterson, Jeremy

APPLICANT: Tetteijn, Herve
APPLICANT: Venter, J. Craig
APPLICANT: Maignani, Vega
APPLICANT: Galeotti, Cesira
APPLICANT: Mora, Manroa
APPLICANT: Ratti, Giulio
APPLICANT: Scarselli, Maria
APPLICANT: Scarlatto, Vincenzo
APPLICANT: Rappuoli, Rino
APPLICANT: Pizze, Mariagrazia
APPLICANT: Grandi, Guido
TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
FILE REFERENCE: 002441.00090
CURRENT APPLICATION NUMBER: US/10/915,740A
CURRENT FILING DATE: 2004-08-11
PRIOR APPLICATION NUMBER: 09/806,866
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: USSN 60/103,794
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: USSN 60/132,068
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: PCT/US99/25373
PRIOR FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 1068
SOFTWARE: PatentIn version 3.2
SEQ ID NO 49
LENGTH: 65632
TYPE: DNA
ORGANISM: Neisseria meningitidis
US-10-915-740A-49

Query Match 70.7%; Score 29; DB 9; Length 65632;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGAGAAACAAATATTGCGGTATAC 41
DB 55054 TATGAGAAACAAATATTGCGGTATAC 55082

RESULT 5
US-10-915-740A-1068/c
Sequence 1068, Application US/10915740A
Publication No. US20050191316A1
GENERAL INFORMATION:
APPLICANT: Frazer, Claire M.
APPLICANT: Hickey, Erin
APPLICANT: Peterson, Jeremy
APPLICANT: Tetteijn, Herve
APPLICANT: Venter, J. Craig
APPLICANT: Maignani, Vega
APPLICANT: Galeotti, Cesira
APPLICANT: Mora, Manroa
APPLICANT: Ratti, Giulio
APPLICANT: Scarselli, Maria
APPLICANT: Scarlatto, Vincenzo
APPLICANT: Rappuoli, Rino
APPLICANT: Pizze, Mariagrazia
APPLICANT: Grandi, Guido
TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
FILE REFERENCE: 002441.00090
CURRENT APPLICATION NUMBER: US/10/915,740A
CURRENT FILING DATE: 2004-08-11
PRIOR APPLICATION NUMBER: 09/806,866
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: USSN 60/103,794
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: USSN 60/132,068
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: PCT/US99/25373
PRIOR FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 1068
SOFTWARE: PatentIn version 3.2

SEQ ID NO 1068
LENGTH: 2242716
TYPE: DNA
ORGANISM: Neisseria meningitidis
US-10-915-740A-1068

Query Match 70.7%; Score 29; DB 9; Length 2242716;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TATGAAAAACAATATTGCGGTATAC 41
DB 79104 TATGAAAAACAATATTGCGGTATAC 79076

RESULT 6
US-09-816-028A-46
Sequence 46, Application US/09816028A
Patent No. US20020042369A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
FILE REFERENCE: 019633-000111US
CURRENT APPLICATION NUMBER: US/09/816,028A
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/118,213
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: US 09/495,406
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 46
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-09-816-028A-46

Query Match 62.0%; Score 25.4; DB 3; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35

RESULT 7
US-10-303-161-46
Sequence 46, Application US/10303161
Publication No. US20030148459A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
FILE REFERENCE: 019633-000111US
CURRENT APPLICATION NUMBER: US/10/303,161
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US/09/816,028
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/118,213
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: US 09/495,406
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 46
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-10-303-161-46

Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35

RESULT 8
US-10-303-118-46
Sequence 46, Application US/10303118
Publication No. US20030157655A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
FILE REFERENCE: 019633-000111US
CURRENT APPLICATION NUMBER: US/10/303,118
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US/09/816,028
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/118,213
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: US 09/495,406
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 46
LENGTH: 41
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-10-303-118-46

Query Match 62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAACAATATTGCGG 35

RESULT 9
US-10-303-128-46
Sequence 46, Application US/10303128
Publication No. US20030157656A1
GENERAL INFORMATION:
APPLICANT: Gilbert, Michel
APPLICANT: Wakarchuk, Warren W.
TITLE OF INVENTION: National Research Council of Canada
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
FILE REFERENCE: 019633-000111US
CURRENT APPLICATION NUMBER: US/10/303,128
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US/09/816,028
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/118,213

```
/ PRIOR FILING DATE: 1999-02-01
/ PRIOR APPLICATION NUMBER: US 09/495,406
/ PRIOR FILING DATE: 2000-01-31
/ NUMBER OF SEQ ID NOS: 49
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 46
/ LENGTH: 41
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
US-10-303-128-46
```

```
Query Match          62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
DB      1 CTTAGAGGTCATATGAAAAAGTTATTATTCCTG 35
```

```
RESULT 10
US-10-303-134-46
/ Sequence 46, Application US/10303134
/ Publication No. US20030157657A1
/ GENERAL INFORMATION:
/ APPLICANT: Gilbert, Michel
/ APPLICANT: Wakarchuk, Warren W.
/ TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
/ TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
/ FILE REFERENCE: 019633-000111US
/ CURRENT FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US/09/816,028
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: US 60/118,213
/ PRIOR FILING DATE: 1999-02-01
/ PRIOR APPLICATION NUMBER: US 09/495,406
/ NUMBER OF SEQ ID NOS: 49
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 46
/ LENGTH: 41
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
US-10-303-134-46
```

```
Query Match          62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
DB      1 CTTAGAGGTCATATGAAAAAGTTATTATTCCTG 35
```

```
RESULT 11
US-10-303-162-46
/ Sequence 46, Application US/10303162
/ Publication No. US20030157658A1
/ GENERAL INFORMATION:
/ APPLICANT: Gilbert, Michel
/ APPLICANT: Wakarchuk, Warren W.
/ TITLE OF INVENTION: National Research Council of Canada
/ TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
/ FILE REFERENCE: 019633-000111US
```

```
/ CURRENT APPLICATION NUMBER: US/10/303,162
/ CURRENT FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US/09/816,028
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: US 60/118,213
/ PRIOR FILING DATE: 1999-02-01
/ PRIOR APPLICATION NUMBER: US 09/495,406
/ NUMBER OF SEQ ID NOS: 49
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 46
/ LENGTH: 41
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
US-10-303-162-46
```

```
Query Match          62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
DB      1 CTTAGAGGTCATATGAAAAAGTTATTATTCCTG 35
```

```
RESULT 12
US-10-735-419-46
/ Sequence 46, Application US/10735419
/ Publication No. US20040180406A1
/ GENERAL INFORMATION:
/ APPLICANT: Gilbert, Michel
/ APPLICANT: Wakarchuk, Warren W.
/ TITLE OF INVENTION: National Research Council of Canada
/ TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
/ FILE REFERENCE: 019633-000111US
/ CURRENT FILING DATE: 2003-12-11
/ PRIOR APPLICATION NUMBER: US/09/816,028A
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: US 60/118,213
/ PRIOR FILING DATE: 1999-02-01
/ PRIOR APPLICATION NUMBER: US 09/495,406
/ NUMBER OF SEQ ID NOS: 49
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 46
/ LENGTH: 41
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:CU-131 5'
US-10-735-419-46
```

```
Query Match          62.0%; Score 25.4; DB 6; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 CTTAGAGGTCATATGAAAAACAATAATTGCGG 35
DB      1 CTTAGAGGTCATATGAAAAAGTTATTATTCCTG 35
```

```
RESULT 13
US-10-820-536-46
/ Sequence 46, Application US/10820536
/ Publication No. US20040203103A1
/ GENERAL INFORMATION:
/ APPLICANT: Gilbert, Michel
```



```

; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/820,536
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
; OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-10-820-536-46

```

```

Query Match      62.0%; Score 25.4; DB 8; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 CTTAGAGGTCATATGAAAAACAATAATTTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAAGTTATTTGCTG 35

```

```

RESULT 14
US-10-845-408-46
; Sequence 46, Application US/10845408
; Publication No. US20040203112A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,408
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
; OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-10-845-408-46

```

```

Query Match      62.0%; Score 25.4; DB 8; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 CTTAGAGGTCATATGAAAAACAATAATTTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAAGTTATTTGCTG 35

```

```

RESULT 15
US-10-845-412-46
; Sequence 46, Application US/10845412
; Publication No. US20040203113A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,412
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C1-131 5'
; OTHER INFORMATION: primer used to amplify and clone ORF 7a
US-10-845-412-46

```

```

Query Match      62.0%; Score 25.4; DB 8; Length 41;
Best Local Similarity 82.9%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 CTTAGAGGTCATATGAAAAACAATAATTTGCGG 35
DB 1 CTTAGAGGTCATATGAAAAAGTTATTTGCTG 35

```

Search completed: April 7, 2006, 18:56:37
Job time : 70.1969 secs

THIS PAGE BLANK (USPTO)


```
/ Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 723641
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-723641

Query Match          55.6%; Score 22.8; DB 6; Length 1095;
Best Local Similarity 92.3%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 AGGTCATATGAAAAACAAATATTG 32
DB 435 ACGGCATATGAAAAACAAATATTG 410

RESULT 3
US-09-925-065A-723642/C
; Sequence 723642, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 723642
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-723642

Query Match          55.6%; Score 22.8; DB 6; Length 1095;
Best Local Similarity 92.3%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
RESULT 4
US-09-925-065A-723643/C
; Sequence 723643, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 723643
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-723643

Query Match          55.6%; Score 22.8; DB 6; Length 1095;
Best Local Similarity 92.3%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 AGGTCATATGAAAAACAAATATTG 32
DB 435 ACGGCATATGAAAAACAAATATTG 410

RESULT 5
US-11-121-086-83
; Sequence 83, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: NIELSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patencin version 3.3
; SEQ ID NO 83
; LENGTH: 187745
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-83

Query Match          55.1%; Score 22.6; DB 14; Length 187745;
Best Local Similarity 86.2%; Pred. No. 7.9e+02;
Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 GGTGATATGAAAAACAAATATTGCGGT 36
DB 66722 GGCACATGATGATTAACAAATATTGCGT 66750

RESULT 6
US-09-925-065A-272562
; Sequence 272562, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
```

Query Match	54.6%; Score 22.4; DB 6;	Length 528;
Best Local Similarity	72.5%; Pred. No. 1.4e+02;	
Matches	29; Conservative 0; Mismatches 11;	Indels 0; Gaps 0;
CY	1 CTTAGAGCTCATATGAAAAACAATAATTCCGTATATA 40 	
Ddb	322 CTTAGGAGTCAATCTCATAACACAAATTAAGCCTTAGA 361 	
RESULT 8		
US-09-925-065A-272564		

```

Query Match          54.6%; Score 22.4; DB 10; Length 533;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1  CTTAGAGGTCAATATGAAAAACAAATAATTTGGCGTTATA 40
          |||||  |||||  |||  |||  |||||  |||  |||
Db       327 CTTAGGAGATCAATATCTCAAAATCACAAATATTAGCCTTAGA 366
          |||||  |||||  |||  |||  |||||  |||  |||

RESULT 10
US-10-301-480-349857
; Sequence 349857, Application US/10301480
; Publication No. US20060057564A1
GENERAL INFORMATION:

```

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349857
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-349857

Query Match      54.6%; Score 22.4; DB 10; Length 533;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGAAAAACAATAATTTGCGGTATA 40
DB      327 CTTAGAGATCATATCTCAATCACAATATTAGCCTTAGA 366

RESULT 11
US-10-301-480-349858
; Sequence 349858, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349858
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-349858

Query Match      54.6%; Score 22.4; DB 10; Length 533;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGAAAAACAATAATTTGCGGTATA 40
DB      327 CTTAGAGATCATATCTCAATCACAATATTAGCCTTAGA 366

RESULT 12
US-10-301-480-963265
; Sequence 963265, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
```

```

; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 963265
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-963265

Query Match      54.6%; Score 22.4; DB 10; Length 533;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGAAAAACAATAATTTGCGGTATA 40
DB      327 CTTAGAGATCATATCTCAATCACAATATTAGCCTTAGA 366

RESULT 13
US-10-301-480-963266
; Sequence 963266, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 963266
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-963266

Query Match      54.6%; Score 22.4; DB 10; Length 533;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1 CTTAGAGGTCATATGAAAAACAATAATTTGCGGTATA 40
DB      327 CTTAGAGATCATATCTCAATCACAATATTAGCCTTAGA 366

RESULT 14
US-10-301-480-963267
; Sequence 963267, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 963267
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-963267
```

Query Match 54.6%; Score 22.4; DB 10; Length 533;
 Best Local Similarity 72.5%; Pred. No. 1.4e+02;
 Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CTTAGAGGTCATATGGAACCAATATTGCGTTATA 40
 |||||
 DB 327 CTTAGAGGTCATATCTCAATATCAATATTAGCCTTAGA 366

RESULT 15

US-09-925-065A-879344
 / Sequence 879344, Application US/09925065A
 / Publication No. US20040181048A1
 / GENERAL INFORMATION:
 / APPLICANT: Wang, David G.
 / TITLE OF INVENTION: Identification and Mapping of Single
 / TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
 / FILE REFERENCE: 108827.135
 / CURRENT APPLICATION NUMBER: US/09/925,065A
 / PRIOR FILING DATE: 2001-08-08
 / PRIOR APPLICATION NUMBER: US 60/243,096
 / PRIOR FILING DATE: 2000-10-24
 / PRIOR APPLICATION NUMBER: US 60/252,147
 / PRIOR FILING DATE: 2000-11-20
 / PRIOR APPLICATION NUMBER: US 60/250,092
 / PRIOR FILING DATE: 2000-11-30
 / PRIOR APPLICATION NUMBER: US 60/261,766
 / PRIOR FILING DATE: 2001-01-16
 / PRIOR APPLICATION NUMBER: US 60/289,846
 / PRIOR FILING DATE: 2001-05-09
 / NUMBER OF SEQ ID NOS: 957086
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 879344
 / LENGTH: 575
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / US-09-925-065A-879344

Query Match 54.6%; Score 22.4; DB 6; Length 575;
 Best Local Similarity 72.5%; Pred. No. 1.5e+02;
 Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 TTAGAGGTCATATGGAACCAATATTGCGTTATAC 41
 |||||
 DB 223 TTAGAGGTCATATGGAATGGAATTAATATTGTTGATTTC 262

Search completed: April 7, 2006, 21:23:54
 Job time : 48.8632 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 15:58:52 ; Search time 307.615 Seconds
(without alignments)
8315.447 Million cell updates/sec

Title: US-09-211-691-4

Perfect score: 45
Sequence: 1 cgacagaaatccgcacacgc.....tttgttaagaatgttttc 45

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Lifting first 45 summaries

Database :

GenEmbl:*
1: gb.ba:*
2: gb.in:*
3: gb.env:*
4: gb.om:*
5: gb.ov:*
6: gb.pac:*
7: gb.ph:*
8: gb.pr:*
9: gb.ro:*
10: gb.ste:*
11: gb.sy:*
12: gb.un:*
13: gb.vl:*
14: gb.hcg:*
15: gb.pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	27	60.0	687	1	NM060146
C 2	27	60.0	1875	1	U04328
C 3	27	60.0	4176	1	NM85YN
C 4	27	60.0	4908	1	NM95053
C 5	27	60.0	110000	1	AE002098
C 6	27	60.0	349980	6	AX043922
C 7	25.8	57.3	239048	14	AC094874
C 8	25	55.6	2603	8	BC016953
C 9	24.6	54.7	152506	9	AC145589
C 10	24.6	54.7	166529	9	AC147621
C 11	24.2	53.8	59262	14	AC107117
C 12	24.2	53.8	185020	14	AC135125
C 13	24	53.3	367	6	COS05083
C 14	24	53.3	237515	14	AC126058
C 15	24	52.9	174372	8	AC099051
C 16	23.8	52.9	217409	8	AC123901
C 17	23.8	52.9	227693	14	AC095313
C 18					

C 19	23.8	52.9	233886	14	AC128261	AC128261 Rattus no
C 20	23.8	52.9	239030	14	AC135695	AC135695 Rattus no
C 21	23.8	52.9	271667	14	AC160572	AC160572 Bos tauru
C 22	23.8	52.9	318143	14	AC095005	AC095005 Rattus no
C 23	23.6	52.4	31	6	CQ895926	CQ895926 Sequence
C 24	23.6	52.4	1174	1	MC150	Z33105 M.capricolu
C 25	23.6	52.4	257159	14	AC094460	AC094460 Rattus no
C 26	23.6	52.4	306172	14	AC103326	AC103326 Rattus no
C 27	23.4	52.0	351	6	CQ460253	CQ460253 Sequence
C 28	23.4	52.0	943	2	AF508884	AF508884 Xyleborus
C 29	23.4	52.0	2759	15	AY436553	AY436553 Theilungl
C 30	23.4	52.0	3812	6	CQ850365	CQ850365 Sequence
C 31	23.4	52.0	3812	8	AK127490	AK127490 Homo sapi
C 32	23.4	52.0	4470	6	CQ841515	CQ841515 Sequence
C 33	23.4	52.0	4470	8	AK124570	AK124570 Homo sapi
C 34	23.4	52.0	55151	15	AB020751	AB020751 Arabidops
C 35	23.4	52.0	110000	14	AC153053	Continuation (2 of
C 36	23.4	52.0	183827	8	AC005899	AC005899 Homo sapi
C 37	23.4	52.0	184841	14	AC129486	AC129486 Homo sapi
C 38	23.2	51.6	720	6	AR397330	AR397330 Sequence
C 39	23.2	51.6	132948	8	HS349A12	AL033520 Human DNA
C 40	23.2	51.6	148259	9	AC104059	AC104059 Mus muscu
C 41	23.2	51.6	175614	14	AC159570	AC159570 Callithri
C 42	23.2	51.6	181669	9	AL645903	AL645903 Mouse DNA
C 43	23.2	51.6	205647	9	AC111132	AC111132 Mus muscu
C 44	23.2	51.6	228380	14	AC133352	AC133352 Rattus no
C 45	23.2	51.6	244451	14	AC125765	AC125765 Rattus no

ALIGNMENTS

RESULT 1
LOCUS NM060146/c 687 bp DNA linear BCT 11-JUN-1997
DEFINITION Neisseria meningitidis CMP-sialic acid synthetase gene, complete cds.
ACCESSION U60146
VERSION U60146.1 GI:1549340
KEYWORDS
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
REFERENCE 1 (bases 1 to 687)
AUTHORS Gilbert,M., Watson,D.C. and Makarchuk,W.W.
TITLE Purification and characterization of the recombinant CMP-sialic acid synthetase from Neisseria meningitidis
JOURNAL Biochemol. Lett. 19, 417-420 (1997)
RECORD 2 (bases 1 to 687)
AUTHORS Gilbert,M., Watson,D.C. and Makarchuk,W.W.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-1996) Institute for Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A 0R6, Canada

FEATURES
source
CDS
1..687
location/Qualifiers
1..687
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="406Y; NRC04030"
/db_xref="taxon:487"
/note="capsule type: Y; LPS type: L3"
1..687
/note="First 40 residues of the N-terminus of the recombinant product were determined experimentally."
/evidence=experimental
/transl_table=11
/product="CMP-sialic acid synthetase"
/protein_id="AAB60780.1"
/db_xref="GI:1549341"
/translation="MEKONIAVILARQNSKGLPKNLKRNQGISLGHITNAISSKC
FDRIIVTDGGLIAEAKNPFVEVVLRAEALASPTASISGVTHALETIGNSGVTL

ORIGIN

LQPTSPLEGTGATGATTAAGATGTTTC 45
LEOPROOLPOAFRNGAIYINDTASLIANNCFIAPTCLYIMSHODSIDIDTELDLQO
AENIINHES"

Query Match 60.0%; Score 27; DB 1; Length 687;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GCTTCTGATTAAGATGTTTC 45
Db 684 GCTTCTGATTAAGATGTTTC 658

RESULT 2 1875 bp DNA linear BCT 15-APR-1994
U04328/c

DEFINITION Neisseria meningitidis NMB CMP-N-acetylneuraminic acid synthetase
(synB) and synX (synX) genes, complete cds.

ACCESSION U04328
VERSION U04328.1 GI:460144

KEYWORDS

Neisseria meningitidis
Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.

1 (bases 1 to 1875)

Swartley, J.S. and Stephens, D.S.
Identification of a genetic locus involved in the biosynthesis of
N-acetyl-D-mannosamine, a precursor of the (alpha 2-->8)-linked
polysialic acid capsule of serogroup B Neisseria meningitidis

J. Bacteriol. 176 (5), 1530-1534 (1994)
8113198

2 (bases 1 to 1875)

Stephens, D.S.
Direct Submission

Submitted (13-DEC-1993) David S. Stephens, Emory University,
Department of Medicine, 69 Butler Street, Atlanta, GA 30303, USA

Location/Qualifiers
1. 1875

/organism="Neisseria meningitidis"
/mol_type="unassigned DNA"

/strain="NMB"
/isolate="M7"

/db_xref="taxon:487"
/note="serogroup B"

52. 1185
/gene="synX"

52. 1185
/gene="synX"
/function="proposed to be involved in the biosynthesis of
N-acetyl-D-mannosamine"

/codon_start=1
/evidence=experimental
/transl_table=11

/product="SynX"
/protein_id="AAI17654.1"

/db_xref="GI:460145"
/translation="MKRIICITGRADFGKLPDLAYIENHDLBIHLIVTGMMNMT
YGRYKEVRENOHTYLFNSIOGEPGAVLGNITTFISRLSDEIEPDVMIHGRLL
BALAGAAGALSSRLVCHIRGELSGVDISIRHSISKSHILVANEOAVRIYVOMG
EKRKHIITGSPDLDMASSTLPSEVEYEGGLPPEYNGISMFHYTTEAHIMPOYA
AQYFKALELGGONITISYPNNDGTESIIQELKQSDKFIAPPSIRFEYFVYLKHA
KFWGNSSAGIRBAPLYGPSIDVGRQSRNMGKSIHTDYETKNIIPDAIQOACSLG
KFEADDFNGCDRTSTERPAEVIINPDTWNSAOKRFIDLN"

1189. 1875
/gene="synB"

1189. 1875
/gene="synB"

/standard_name="CMP-NANA synthetase"
/codon_start=1
/transl_table=11

/product="CMP-N-acetylneuraminic acid synthetase"

gene
CDS

gene
CDS

ORIGIN

/protein_id="AAI17655.1"
/db_xref="GI:460146"
/translation="MEKONTAVILIAARONSKLPKXNRYKNGISLIGHTTAAIASSKC
FDRITVSDGSLAEKONRGVEYVLRPAELASTPSSISGVHAELETIGNSCTVTL
LQPTSPLEGTGATGATTAAGATGTTTC 45
LEOPROOLPOAFRNGAIYINDTASLIANNCFIAPTCLYIMSHODSIDIDTELDLQO
AENIINHES"

Query Match 60.0%; Score 27; DB 1; Length 1875;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GCTTCTGATTAAGATGTTTC 45
Db 1872 GCTTCTGATTAAGATGTTTC 1846

RESULT 3 4176 bp DNA linear BCT 18-APR-2005
NMBSYN/c

DEFINITION N meningitidis (group B) ctrA, synA, synB, synC and
sialyltransferase genes.

ACCESSION X78068
VERSION X78068.1 GI:530037

KEYWORDS CMP-NeuNAc synthetase; ctrA gene; sialyltransferase; synA gene;
synB gene; synC gene.

1 (bases 1 to 4176)

Neisseria meningitidis serogroup B
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.

1 (bases 1 to 4176)

Ganguli, S., Zapata, G., Wallis, T., Reid, C., Boulnois, G., Vann, W.F.
and Roberts, I.S.

Molecular cloning and analysis of genes for sialic acid synthesis
in Neisseria meningitidis group B and purification of the
meningococcal CMP-NeuNAc synthetase enzyme

J. Bacteriol. 176 (15), 4583-4589 (1994)
8045888

2 (bases 1 to 4176)

Ganguli, S.
Direct Submission

Submitted (09-MAR-1994) S. Ganguli, University of Leicester, Dept
of Microbiology, PO Box 138, Medical Sciences Building, University
Road, Leicester LE1 9HN, UK

Related sequences: M57677 & M64389.
Location/Qualifiers

1. 4176

/organism="Neisseria meningitidis serogroup B"
/mol_type="genomic DNA"

/isolate="NCTC8249"
/db_xref="taxon:491"

complement(<1. 657)
/gene="ctrA"

complement(<1. 657)
/gene="ctrA"

792. 1925
/gene="synA"

792. 1925
/gene="synA"

/codon_start=1
/transl_table=11

/protein_id="CAA54982.1"

/db_xref="GI:530038"

/db_xref="GOA:Q57141"

/db_xref="InterPro:IPR003331"

/db_xref="UniProt/TREMBL:Q57141"

/translation="MKRIICITGRADFGKLPDLAYIENHDLBIHLIVTGMMNMT
YGRYKEVRENOHTYLFNSIOGEPGAVLGNITTFISRLSDEIEPDVMIHGRLL
BALAGAAGALSSRLVCHIRGELSGVDISIRHSISKSHILVANEOAVRIYVOMG
EKRKHIITGSPDLDMASSTLPSEVEYEGGLPPEYNGISMFHYTTEAHIMPOYA
AQYFKALELGGONITISYPNNDGTESIIQELKQSDKFIAPPSIRFEYFVYLKHA
KFWGNSSAGIRBAPLYGPSIDVGRQSRNMGKSIHTDYETKNIIPDAIQOACSLG
KFEADDFNGCDRTSTERPAEVIINPDTWNSAOKRFIDLN"

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
LOCUS NM95053/c	Neisseria meningitidis sial (sial), CMP-NeuNac synthetase (sialb),	M95053	M64289	M95562	M95053.1	GI:520732
DEFINITION	sialc (sialc), and siald (siald) genes, complete cds.					
ACCESSION						
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						

REFERENCE	Bacteria: Proteobacteria; Betaproteobacteria; Neisseriales;
AUTHORS	Neisseriaceae; Neisseria.
TITLE	1 (bases 2861 to 4641) Frosch,M., Edwards,U., Bousset,K., Krause,B. and Weisgerber,C. Evidence for a common molecular origin of the capsule gene loci in gram-negative bacteria expressing group II capsular polysaccharides Mol. Microbiol. 5 (5), 1251-1263 (1991)
JOURNAL	1659649
PUBMED	2 (bases 1210 to 1987) Edwards,U. and Frosch,M. Sequence and functional analysis of the cloned Neisseria meningitidis CMP-NeuNac synthetase FEBS Microbiol. Lett. 75 (2-3), 161-166 (1992)
REFERENCE	1398032
AUTHORS	3 (bases 1 to 4908) Edwards,U., Muller,A., Hammerschmidt,S. and Frosch,M. Molecular analysis of the biosynthesis pathway of the alpha 2,8 polysialic acid capsule by Neisseria meningitidis serogroup B unpublished
JOURNAL	4 (bases 1210 to 1987)
REFERENCE	Frosch,M. Direct Submission Submitted (28-Oct-1992) Matthias Frosch, Medizinische Hochschule Hannover, Institut fur Medizinische Mikrobiologie, Konstanty Gutschow Strasse 8, Hannover, Germany, 30623
AUTHORS	5 (bases 1 to 4908) Frosch,M. Direct Submission Submitted (25-MAY-1994) Matthias Frosch, Medizinische Hochschule Hannover, Institut fur Medizinische Mikrobiologie, Konstanty Gutschow Strasse 8, Hannover, Germany, 30623
JOURNAL	On or before Apr 4, 2002 this sequence version replaced gi:150354, gi:497288.
COMMENT	Location/Qualifiers
FEATURES	1..4908 /organism="Neisseria meningitidis" /mol_type="genomic DNA" /isolate="B1940" /species.host="Homo sapiens" /db_xref="taxon:487"
SOURCE	51..56 /standard_name="-10 signal"
	174..1307 /gene="siaA"
gene	174..1307 /gene="siaA"
CDS	174..1307 /gene="siaA" /codon_start=1 /tranai_table=11 /product="Sia"
-10_signal	/protein_id="AAA20475.1" /db_xref="GI:520733"
gene	/translation="MKRIILCTGTGRADFGKLKPLAYIENHPDLLEHLIVTGAMMMKT YGVTVGAVRENYOHTYLPSNOICGBPMGAIVAGNTTETLSRSDLETSPDMVTHGRL BALAGAVALSSRLVCHIEGGELSGTVDSIRHSISKSHITLVANEQAVITLVONG EKRRKHIIIGSPLDVWASSTPLSLBEVKEYGLPYENVGISNHPTVTFAHLPDYA AQCFKALESIGNDIISIPNNIDGESIIQLLKYOSDKFIAPSRIFBYFLVLTKIA KFWNGWSAGIREAPLYGPSIDVGTRONNRHGKSIITHDYEKKVI PDAIQACSGISG KPFADDTENGCDIRTERAFAYINNPETWNSAQKRFDIND."
CDS	1311..1997 /gene="siab"
gene	1311..1997 /gene="siab"
CDS	1311..1997 /gene="siab" /citation=[2] /codon_start=1 /tranal_table=11 /product="CMP-NeuNac synthetase" /protein_id="AAA20476.1"
CDS	/db_xref="GI:520734" /translation="MEKONIATVILARONSGLPDKNIARKNGISILKHTINAIASKC PDILIYSOGGLIAEBAKNRGVVVARPAELASDTASISGTHALETTIGNSGCVTL LGTPSPRTCAHI REAFSLDEKIKSYVSACMEHRPLTKLIQTINGRARAPRHRS LEDPKOOPAEFRPNGALITYNDIASLIANNCFPIATPKIYNHSDSIDIDLELDQQ AENILNHNKE"

gene 1998. 3047
/gene="siac"
CDS 1998. 3047
/gene="siac"
/citation=[1]
/codon_start=1
/transl_table=1
/product="siac"
/protein_id="AAA20477.1"
/db_xref="GI:520735"
/translation="MNNNEPFIKGRSVGNHEPLICEIGINHGSLKTAPEWDA
VYAGAEVGHOTIIVEDWEDAKOVIIPENADVSIEIMERCALNEEDETIKKEYES
KMIFISPPSPAAALRIORMDIPAKIGSGCANNPIKIVASPGKPIITSTGANSI
ESIKKSVETIRACVAPYALHCTNITPYPEYEVRRGAMDLSEAPDAITGSDITLD
NTACGVALGSGISIERHTDMDRPGPIVCSMPDITKEIKQGHAKLKGKGD
TIAEKPDPAPASVADKDIKKEGELLGNNLWVRKNGDFSVNEYETLFGVAA
CNIRGAQIKKTDIE"
3131. 4618
/gene="siad"
3131. 4618
/gene="siad"
/citation=[3]
/codon_start=1
/transl_table=1
/db_xref="GI:520736"
/protein_id="AAA20478.1"
/translation="MIKKIKKALFQPKKFFQDSMWLTTPSYLTTPRNNLPIYSNLQ
LNQVSLIKIQKLTNNLVIITTSKMLPKLVHQSANKLPSYITLPELPSNNIT
PKKLIYRSYKKNINIQPAHYNLSTFGHYSYLSIAKKNIITTHLIDESTGYAP
LSESPHPTKLERLYIGNNLHKGIDHFDLHVPFPYAKKIPAKKNRFPYAG
GISINNNIANKKOISKNDYIFVSQRPISDDLKYSIVLINSISIQIGKIFIK
LHPKKNYVMSLFNWEINPRLVINEPPLLEPIEPLIYLTNPQIGILASSLIYT
PLISSTOCLSTGELLINLLOKYSWENTEMIOEHLEIKKFNPTIINDLNGVSNP
LKTREPTETIKSAEFAKSKYQAIPTYNOLASNNITLGHAKLYTNALYNVQ
LYKMSYDIFYDNISVDFHSKOKLTWEIKYISADNRIGDR"

ORIGIN

Query Match 60.0%; Score 27; DB 1; Length 4908;
Best local similarity 100.0%; Pred. No. 8.9;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 19 GCTTTCCTGTGATTAAAGATGTTTC 45
Db 1994 GCTTTCCTGTGATTAAAGATGTTTC 1968

RESULT 5
AE002098_00
WPCOMMENT

Sequence split into 23 fragments LOCUS AE002098 Accession AE002098
Fragment Name Begin End
AE002098_00 1 110000
AE002098_01 100001 210000
AE002098_02 200001 310000
AE002098_03 300001 410000
AE002098_04 400001 510000
AE002098_05 500001 610000
AE002098_06 600001 710000
AE002098_07 700001 810000
AE002098_08 800001 910000
AE002098_09 900001 1010000
AE002098_10 1000001 1110000
AE002098_11 1100001 1210000
AE002098_12 1200001 1310000
AE002098_13 1300001 1410000
AE002098_14 1400001 1510000
AE002098_15 1500001 1610000
AE002098_16 1600001 1710000
AE002098_17 1700001 1810000
AE002098_18 1800001 1910000
AE002098_19 1900001 2010000
AE002098_20 2000001 2110000
AE002098_21 2100001 2210000

AE002098_22 2200001 2272360
LOCUS AE002098 2272360 bp DNA circular BCT 26-MAY-2005
DEFINITION Neisseria meningitidis MC58, complete genome.
ACCESSION AE002098 AE002359-AE002564
VERSION AE002098.2 GI:66731897
KEYWORDS

SOURCE
ORGANISM

Neisseria meningitidis MC58
Neisseria meningitidis MC58
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;

REFERENCE

1 (bases 1 to 2272360)
Tetzelin H., Saunders N.J., Heidelberg J., Jeffries A.C.,
Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
Dodson R.V., Nelson M.C., Grimm M.L., DeBoy R., Peterson J.D.,
Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
Douglass B.A., Mason T., Ciecko A., Parksey D.S., Blair E.,
Cittone H., Clark E.B., Cotton M.D., Uterback T.R., Khouri H.,
Qin H., Vamathevan J., Gill J., Scariato V., Maignani V.,
Piazza M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R.,
Rappuoli R. and Venter J.C.
Complete genome sequence of Neisseria meningitidis serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)
2 (bases 1 to 2272360)
Tetzelin H., Saunders N.J., Heidelberg J., Jeffries A.C.,
Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Ciecko A.,
Pedon J.F., Dodson R.V., Nelson W.C., Grimm M.L., Peterson J.D.,
Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
Douglass B.A., Mason T.M., Parksey D.S., Blair E., Clitone H.,
Clark E.B., Cotton M.D., Uterback T.R., Khouri H.M., Qin H.,
Vamathevan J., Gill J., Scariato V., Maignani V., DeBoy R.T.,
Piazza M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R.,
Rappuoli R. and Venter J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
3 (bases 1 to 2272360)
Tetzelin H., Saunders N.J., Heidelberg J., Jeffries A.C.,
Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Ciecko A.,
Pedon J.F., Dodson R.V., Nelson W.C., Grimm M.L., Peterson J.D.,
Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
Douglass B.A., Mason T.M., Parksey D.S., Blair E., Clitone H.,
Clark E.B., Cotton M.D., Uterback T.R., Khouri H.M., Qin H.,
Vamathevan J., Gill J., Scariato V., Maignani V., DeBoy R.T.,
Piazza M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R.,
Rappuoli R. and Venter J.C.
Direct Submission
Submitted (18-MAY-2005) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT
REMARK
On or before May 26, 2005 this sequence version replaced
Sequence update by submitter

gi:7225225, gi:7225241, gi:7225245, gi:7225252, gi:7225254,
gi:7225269, gi:7225277, gi:7413418, gi:7413419, gi:7225303,
gi:7225314, gi:7225327, gi:7225337, gi:7413420, gi:7413421,
gi:7413423, gi:7225394, gi:7225407, gi:7225416, gi:7225428,
gi:7225435, gi:7225442, gi:7225455, gi:7225470, gi:7225484,
gi:7225498, gi:7225506, gi:7225512, gi:7225523, gi:7225529,
gi:7225537, gi:7225550, gi:7225561, gi:7225573, gi:7225586,
gi:7413428, gi:7413429, gi:7225631, gi:7225640,
gi:7225650, gi:7225659, gi:7225669, gi:7225683, gi:7225688,
gi:7225697, gi:7413430, gi:7413433, gi:7225726, gi:7225740,
gi:7225757, gi:7225766, gi:7225776, gi:7225783, gi:7225793,
gi:7225809, gi:7225816, gi:7225823, gi:7225845, gi:7225853,
gi:7225863, gi:7225876, gi:7225889, gi:7225898, gi:7225913,
gi:7225923, gi:7225930, gi:7225939, gi:7225948, gi:7225959,
gi:7225969, gi:7225986, gi:7225997, gi:7226012, gi:7226023,
gi:7226036, gi:7226049, gi:7226062, gi:7226072, gi:7226083,
gi:7226100, gi:7226125, gi:7226133, gi:7226139, gi:7226149,
gi:7226162, gi:7226173, gi:7226185, gi:7413445, gi:7413446,
gi:7226229, gi:7226238, gi:7226246, gi:7226268, gi:7226281,
gi:7226273, gi:7226282, gi:7226297, gi:7226304, gi:7226311,
gi:7226320, gi:7226335, gi:7226350, gi:7226363, gi:7413449,

Query Match 60.0%; Score 27; DB 6; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GCTTCCTTGATAGAGATGTTTC 45
 |||||
 Db 78420 GCTTCCTTGATAGAGATGTTTC 78446

RESULT 7
 AC094874/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-5N22. *** SEQUENCING IN PROGRESS ***
 7 unordered pieces.
 AC094874
 AC094874.7 GI:30466366
 HTG: HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
 1 (bases 1 to 239048)
 Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D.,
 Anyaledechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Bialwalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Cavazos, I., Caesar, H., Calderon, B.,
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Calderon, B.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Creel, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
 Hollins, B., Howell, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, J., Kovar, C.,
 Kows, C., Kratt, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, D.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenshuwa, L., Loulsegue, H., Lozano, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, R., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mamnney, S., McLeod, M. P., McNeill, T. Z., Meenan, B.,
 Milosavljevic, A., Miner, G., Minja, B., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Muniasa, M., Murphy, M., Naif, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwackelmele, O., Okunolu, G., Olarnunagsoon, A., Pal, S., Parks, K.,
 Pasernak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,
 Plopper, F., Polinder, A., Popovic, D., Primus, E., Pu, L., L.,
 Puato, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reich, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rivers, C., Rodkey, T., Rojas, R., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajls, D.,
 Shee, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, P.,
 Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valae, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,
 Williams, G., Willison, R., Wleczek, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstein, G., and Gibbs, R. A.
 Direct Submission

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Unpublished
 2 (bases 1 to 239048)
 Worley, K. C.
 Direct Submission
 Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 239048)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 9, 2003 this sequence version replaced gi:24818336.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: GBPK
 Center clone name: CH230-5N22

 Summary Statistics
 Assembly program: Atlas;
 Consensus quality: 228277 bases at least Q40
 Consensus quality: 230525 bases at least Q30
 Consensus quality: 232432 bases at least Q20
 Estimated insert size: 234098; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 226899: contig of 226899 bp in length
 * 226900: gap of unknown length
 * 227000: contig of 4914 bp in length
 * 231913: contig of 1140 bp in length
 * 232013: gap of unknown length
 * 232014: contig of 1140 bp in length
 * 232154: gap of unknown length
 * 233253: contig of 1310 bp in length
 * 233254: gap of unknown length
 * 234563: contig of 1473 bp in length
 * 234664: gap of unknown length
 * 236137: contig of 1063 bp in length
 * 236237: gap of unknown length
 * 237299: contig of 1063 bp in length
 * 237300: gap of unknown length
 * 237399: gap of unknown length
 * 239048: contig of 1649 bp in length.
 Location/Qualifiers
 1. 239048
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-5N22"
 1. 1354
 /note="wgs_end_extension"

FEATURES
 source
 misc_feature
 misc_feature

[illegible]

Through the I.M.A.G.E. Consortium/INTL at: <http://image.lnh.gov>
 Series: IRAX Plate: 22 Row: f Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarly but not
 identity to protein.

Location/Qualifiers
1. .2603
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3896398"
/tissue_type="Pancreas", epithelioid carcinoma"
/clone_id="NH MGC_70"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match	55.6%	Score 25	DB 8	Length 2603
Best Local Similarity	75.6%	Pred. 0	56	
Matches 31	Conservative	0	Mismatches 10	Indels 0
				Gaps 0

QY 2 GACAGATTCCGCCACCGCTTCTGTGATTAGAATGTT 42
1613 GACACGATTCGCGCTTCACTGTCTTCTCAGTAACAAGTT 1573
Db

RESULT 9	AC145589/c	LOCUS	AC145589	152506 bp	DNA	linear	ROD 15-MAY-2004
DEFINITION		Mus musculus BAC clone RP4-174M10 from chromosome Y, complete					

ACCESSION	AC145589	
VERSION	AC145589.3	GI:45120384
KEYWORDS	HTG.	
SOURCE	Mus musculus	(house mouse)
ORGANISM	Mus musculus	

REFERENCE
1 (bases 1 to 152506)

REFERENCE
AUTHORS
TITLE
JOURNAL
2. (baaes 1 to 152506)
1. (baaes 1 to 152506)
Vanbrunt A., Van Brunt A., Haglund K., Meyer R. and Haekenson, W
The sequence of Mus musculus BAC clone RP24-174M10
Unpublished (2001)
2. (baaes 1 to 152506)

2 (bases 1 to 152506)
Wilson, R.K.
Direct Submission
Submitted (19-JUN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (22-Oct-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 152506)
Wilson, R.K.
Direct Submission

JOURNAL Submitted (22-Oct-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 152506)
AUTHORS Wilson, R.K.
TITLE Direct Submission
ENTRY Submitted (02-Nov-2003) Genome Sequencing Center, 4444 Forest Park

RECEIVED
JOURNAL
TITLE
DIRECT SUBMISSION
Submitted (05-MAR-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 152506)
Wilson, R.K.
nrcslc Submission
TITLE

TITLE Direct Submission
 JOURNAL Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 COMMENT On Mar 5, 2004 this sequence version replaced gi:37806544.
 ----- Genome Center

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: M05SC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu

Contact: submissions@watson.wustl.edu
----- Summary Statistics -----
Center project name: M_BB0174M10

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The position of this clone was established as part of a collaboration between the Mouse Chromosome Y Mapping Project (Jessica E. Alfoldi, Helen Skalecky, Steve Rozen, and David C. Page) at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

SOURCE INFORMATION:

The RPc1-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC139328.

FEATURES

source

1.152506
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="Y"
/map="Y"
/clone="RP24-174M10"
/clone_1ib="RPc1-24"
1.43
/note="Sequence derived from one plasmid subclone."
8009.8525
/note="Unresolved tandem repeat."
58030.58040
/note="Sequence derived from one plasmid subclone."
81477.81543
/product="tRNA-Ser"
/note="likely pseudogene (HMM Sc=16.26 / Sec struct Sc=4.85)"

ORIGIN

Query Match 54.7%; Score 24.6; DB 9; Length 152506;
Best Local Similarity 76.9%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 30; Conservative 0; Mismatches 9;

6 GAATTCGGCCACCGCTTCTGATGATGATGTTT 44

Db 122822 GAATTCGCCCAATGCTTCCATGTGAACCGAATGTTT 122784

RESULT 10
AC147621/c 166529 bp DNA linear ROD 15-MAY-2004
LOCUS AC147621
DEFINITION Mus musculus BAC clone RP24-315M14 from chromosome Y, complete sequence.
ACCESSION AC147621
VERSION AC147621.3 GI:46358243
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 166529)

AUTHORS

VanBrunt, A., Van Brunt, A., Cotton, M. and Bjelicki, L.

TITLE

The sequence of Mus musculus BAC clone RP24-315M14

REFERENCE

2 (bases 1 to 166529)

AUTHORS

Wilson, R.K.

TITLE

Direct Submission

JOURNAL

Submitted (22-DEC-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE

3 (bases 1 to 166529)

AUTHORS

Wilson, R.K.

TITLE

Direct Submission

JOURNAL

Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE

4 (bases 1 to 166529)

AUTHORS

Wilson, R.K.

TITLE

Direct Submission

JOURNAL

Submitted (10-APR-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE

5 (bases 1 to 166529)

AUTHORS

Wilson, R.K.

TITLE

Direct Submission

JOURNAL

Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE

On Apr 10, 2004 this sequence version replaced gi:40539137.

COMMENT

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The position of this clone was established as part of a collaboration between the Mouse Chromosome Y Mapping Project (Jessica E. Alfoldi, Helen Skalecky, Steve Rozen, and David C. Page) at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

SOURCE INFORMATION:

The RPc1-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC145589 and AC140203.

FEATURES

source

1.166529
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="Y"
/map="Y"
/clone="RP24-315M14"
/clone_1ib="RPc1-24"
271.1335
/rpt_family="ERVK"

repeat_region


```
repeat_region 1656. .1731 /rpt_family="ERVK"
repeat_region 1803. .1902 /rpt_family="ERVK"
repeat_region 3764. .3878 /rpt_family="MaLR"
repeat_region 3900. .4490 /rpt_family="L1"
repeat_region 4489. .6799 /rpt_family="L1"
repeat_region 6850. .7034 /rpt_family="L1"
repeat_region 7030. .7186 /rpt_family="MER1_type"
repeat_region 7187. .7312 /rpt_family="MaLR"
repeat_region 7313. .7451 /rpt_family="Alu"
repeat_region 8534. .8720 /rpt_family="MaLR"
repeat_region 11457. .11659 /rpt_family="B2"
repeat_region 14027. .14118 /rpt_family="MaLR"
repeat_region 18092. .18259 /rpt_family="Alu"
repeat_region 18603. .18652 /rpt_family="L1"
repeat_region 18773. .18839 /rpt_family="MaLR"
repeat_region 18890. .19170 /rpt_family="ERV1"
repeat_region 19462. .19733 /rpt_family="MaLR"
repeat_region 20000. .20241 /rpt_family="ERVK"
repeat_region 20825. .20971 /rpt_family="MaLR"
repeat_region 21706. .21842 /rpt_family="Alu"
repeat_region 22418. .22549 /rpt_family="Alu"
repeat_region 23262. .23397 /rpt_family="Alu"
repeat_region 24072. .24220 /rpt_family="Alu"
repeat_region 24833. .24967 /rpt_family="Alu"
repeat_region 25543. .25694 /rpt_family="Alu"
repeat_region 26172. .26436 /rpt_family="Alu"
repeat_region 26530. .26668 /rpt_family="L1"
repeat_region 27249. .27382 /rpt_family="Alu"
repeat_region 27705. .27975 /rpt_family="Alu"
repeat_region 28014. .28157 /rpt_family="L1"
repeat_region 28807. .28945 /rpt_family="Alu"
repeat_region 29573. .29707 /rpt_family="Alu"
repeat_region 30236. .30295 /rpt_family="Alu"
repeat_region 30374. .30452 /rpt_family="ERV1"
repeat_region 30434. .30487 /rpt_family="Alu"
repeat_region 31110. .31170 /rpt_family="B4"
repeat_region 31866. .31996 /rpt_family="Alu"
```

```
repeat_region 32719. .32848 /rpt_family="Alu"
repeat_region 33198. .33905 /rpt_family="Alu"
repeat_region 35982. .36065 /rpt_family="L1"
repeat_region 36278. .37477 /rpt_family="ERVK"
repeat_region 38048. .38946 /rpt_family="ERVK"
repeat_region 39943. .40253 /rpt_family="ERVK"
repeat_region 40482. .40939 /rpt_family="ERV1"
repeat_region 41385. .41423 /rpt_family="ERVK"
repeat_region 41424. .41534 /rpt_family="U2"
repeat_region 41557. .41670 /rpt_family="L1"
repeat_region 42332. .42519 /rpt_family="L1"
repeat_region 42520. .42578 /rpt_family="L1"
repeat_region 43031. .43320 /rpt_family="B4"
repeat_region 43771. .43925 /rpt_family="L1"
repeat_region 43932. .44030 /rpt_family="B4"
repeat_region 44265. .44588 /rpt_family="L1"
repeat_region 45305. .45539 /rpt_family="MaLR"
repeat_region 45825. .46527 /rpt_family="L1"
repeat_region 46533. .47622 /rpt_family="L1"
repeat_region 47688. .48296 /rpt_family="L1"
repeat_region 48297. .48449 /rpt_family="L1"
```

```
Query Match 54.7%; Score 24.6; DB 9; Length 166529;
Best Local Similarity 76.9%; Pred. No. 1e+02;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 6 GAATTCGGCCAGCGCTTCTTGATAGATGTTT 44
DB 15649 GGATTCGCCCAATGCTTTCATGTGAACTGATGTTT 15611
```

```
RESULT 11
AC107117 3/c
WPCOMMENT
```

```
Sequence split into 4 fragments LOCUS AC107117 Accession AC107117
Fragment Name Begin End
AC107117_0 1 110000
AC107117_1 100001 210000
AC107117_2 200001 310000
AC107117_3 300001 359262
Continuation (4 of 4) of AC107117 from base 300001 (AC107117 Rattus norvegicus clone CH;
```

```
Query Match 53.8%; Score 24.2; DB 14; Length 59262;
Best Local Similarity 71.1%; Pred. No. 1.4e+02;
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 1 CGACGAATTCGGCCAGCGCTTCTTGATAGATGTTTC 45
DB 48182 CTGAGGCTTCACCAACGCTTATGAAATGATGTTTC 48138
```

```
RESULT 12
```

AC135125/c
 LOCUS AC135125 185020 bp DNA linear HTG 20-NOV-2002
 DEFINITION Rattus norvegicus clone CH230-265D13, *** SEQUENCING IN PROGRESS
 ACCESSION AC135125
 VERSION AC135125.2 GI:25138669
 KEYWORDS HTG, HTGS, PHASE1, HTGS_DRAFT, HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
 1 (bases 1 to 185020)
 Muzny,D.,Marie., Metzker,M.,Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biewald,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Buttrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Baccotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshuwa,L., Louisedge,H., Lozano,R.J., Lu,X., Ma,D., Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawliny,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munitasa,M., Murphy,M., Nait,L., Nankervyls,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwackelmehe,O., Okunolu,G., Olampunsgoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.-L., Putzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,R., Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,M., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartbeyn,A., Sisson,I., Sitter,C.D., Smjs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soza,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Uemami,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G., and Gibbs,R.A.

JOURNAL
 COMMENT
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Nov 20, 2002 this sequence version replaced gi:23507714.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KCZ2
 Center clone name: CH230-265D13
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 179552 bases at least Q40
 Consensus quality: 181077 bases at least Q30
 Consensus quality: 181866 bases at least Q20
 Estimated insert size: 181422; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 183656: contig of 183656 bp in length
 * 183657 183756: gap of unknown length
 * 183757 185020: contig of 1264 bp in length.
 Location/Qualifiers
 1..185020
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-265D13"
 1..1163
 /note="wgs end_extension
 clone end:77"
 misc_feature
 3257..4125
 /note="clone_boundary
 clone end:77
 site:
 end_sequence:RXAEC19TY"
 complement(178207..179102)
 /note="clone_boundary
 clone_end:Sp6
 site:
 end_sequence:RXAEC19TYB"
 17976..181250
 /note="wgs end_extension
 clone_end:Sp6"
 misc_feature
 182143..183656
 /note="wgs end_extension
 clone end:Sp6"
 183657..183756
 /estimated_length=unknown
 gap
 ORIGIN

Query Match 53.8%; Score 24.2; DB 14; Length 185020;
 Best Local Similarity 71.1%; Pred. No. 1.5e+02;
 Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CGACGAAATTCGGCCAGCCGCTTCTTGATTAAGAAATGTTTC 45
 DB 177225 CTAGAGGCTTCACACCGCTACTATGAAATGACTGTTCC 177181

RESULT 13
 LOCUS CQ505083 367 bp DNA linear PAT 30-JAN-2004
 DEFINITION Sequence 36950 from Patent WO0160860.
 ACCESSION CQ505083
 VERSION CQ505083.1 GI:41471149
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

REFERENCE
 AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
 TITLE Genes differentially expressed in human prostate cancer and their use
 JOURNAL Patent: WO 0160860-A 36950 23-AUG-2001;
 FEATURES
 SOURCE Millennium Predictive Medicine, Inc. (US)
 Location/Qualifiers
 1..367
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 53.3%; Score 24; DB 6; Length 367;
 Best Local Similarity 75.0%; Pred. No. 1.3e+02;
 Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAGAAATTCGGCCAGCCGCTTCTTGATTAAGAAATGTTT 43
 DB 271 CACAATTAGCCCTCTACTCTTCTTGATTAAGAAATGTTT 310

RESULT 14
 LOCUS CQ510708 367 bp DNA linear PAT 30-JAN-2004
 DEFINITION Sequence 42575 from Patent WO0160860.
 ACCESSION CQ510708
 VERSION CQ510708.1 GI:41476972
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

REFERENCE
 AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
 TITLE Genes differentially expressed in human prostate cancer and their use
 JOURNAL Patent: WO 0160860-A 42575 23-AUG-2001;
 FEATURES
 SOURCE Millennium Predictive Medicine, Inc. (US)
 Location/Qualifiers
 1..367
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 53.3%; Score 24; DB 6; Length 367;
 Best Local Similarity 75.0%; Pred. No. 1.3e+02;
 Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAGAAATTCGGCCAGCCGCTTCTTGATTAAGAAATGTTT 43
 DB 271 CACAATTAGCCCTCTACTCTTCTTGATTAAGAAATGTTT 310

RESULT 15
 LOCUS AC126059 237515 bp DNA linear HTG 13-NOV-2002
 DEFINITION Rattus norvegicus clone CH230-9C16, *** SEQUENCING IN PROGRESS ***
 AC126059
 AC126059
 AC126059.3 GI:24940978
 HTG: HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS Muzny, D., Marie, Metzger, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bisswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Derramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Gladys, M., Guertel, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huylk, S., Hune, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyvet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Labow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshew, L., Loulsegged, H., Lozano, R. J., Lu, X., Ma, T., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaekelemeh, O., Okwunonu, G., Olarunmugbon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Prims, E., Pu, L., Pu, M., Putroz, J., Rachlin, E., Reeves, K., Reiter, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Slason, I., Sitter, C. D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J., Steimler, M., Strong, R., Sutton, A., Svetek, A., Taylor, Z., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Tjorja, P., Umant, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.

TITLE
 JOURNAL Direct Submission
 REFERENCE Unpublished
 AUTHORS 2 (bases 1 to 237515)
 TITLE Worley, K. C.
 Direct Submission

CC in formation of a nucleotide sugar which is a saccharide donor for a
CC glycosyltransferase reaction. The fusion protein is useful in the
CC enzymatic synthesis of oligosaccharides. The fusion proteins are able to
CC catalyze more than one reaction involved in the enzymatic synthesis. This
CC is useful for the development of therapeutic agents that have specific
CC carbohydrate structures. Carbohydrates are involved in recognition
CC elements on the surface of cells. The fusion protein can be used for the
CC synthesis of both natural carbohydrates and synthetic derivatives with
CC novel properties. The fusion polypeptide allows two glycosyltransferase
CC reactions in a single vessel, provides improved yields of end products.
CC Additionally, cleanup and disposal of extra solvents and by-products is
CC reduced. The fusion protein can also use directly different donor
CC analogues and various acceptors with a terminal galactose residue

CC SQ Sequence 45 BP; 10 A; 12 C; 8 G; 15 T; 0 U; 0 Other;

Query Match 100.0%; Score 45; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 6.6e-09;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGACGAAATTCGCGCACCGCTTCTTGATTAAGAAGTTTTC 45
1 CGACGAAATTCGCGCACCGCTTCTTGATTAAGAAGTTTTC 45
Db

RESULT 2
AAA81833
ID AAA81833 standard; DNA; 623 BP.

AC AAA81833;

DT 04-DEC-2000 (first entry)

DE N. meningitidis partial DNA sequence gnm_380 SEQ ID NO:380.

KM Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.

OS Neisseria meningitidis.

PN WO200022430-A2.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-US023573.

PR 09-OCT-1998; 98US-0103794P.

PR 30-APR-1999; 99US-0132068P.

PA (CHIR) CHIRON CORP.

PI Frazer CM, Hickey B, Peterson J, Tettelin H, Venter JC;
PI Massignani V, Galeotti C, Mora M, Ratti G, Scarcelli M, Scarlato V;
PI Rappelli R, Piza M;

DR WPI; 2000-318079/27.

PT Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT in the diagnosis and treatment of N. meningitidis infection and other
PT Neisserial infections, for example, N.gonorrhoea.

PS Claim 7; Page 1611; 1760pp; English.

XX The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AAA81453 to AAA82414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA
CC sequences and their corresponding proteins; AAA81254 to AAA81259 and
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid

CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neisseriae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions

CC SQ Sequence 623 BP; 174 A; 118 C; 135 G; 196 T; 0 U; 0 Other;

Query Match 60.0%; Score 27; DB 3; Length 623;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 GCTTCCCTTGATTAAGAAGTTTTC 45
161 GCTTCCCTTGATTAAGAAGTTTTC 187
Db

RESULT 3
AA224472/C
ID AA224472 standard; DNA; 687 BP.

AC AA224472;

DT 17-FEB-2000 (first entry)

DE N. meningitidis siab DNA.

KM CMP sialylate synthetase; siab; cytidine monophosphate; detection;
KW CMP-activated N-acetylneuraminic acid; sialyl acceptor; ds.

OS Neisseria meningitidis.

PN DE19913206-A1.

PD 07-OCT-1999.

PF 24-MAR-1999; 99DE-01013206.

PR 26-MAR-1998; 98DE-01013426.

PA (FESS/) FESSNER W D.

PI Fesener W, Knorst M;

DR WPI; 2000-024923/03.
P-PsDB; AAY50798.

PT Production of CMP-sialate synthetase protein useful for producing CMP-
PT activated N-acetylneuraminic acid analogs and non-naturally sialylated
PT oligosaccharides and glycoconjugates.

PS Claim 5; Page 19-20; 26pp; German.

XX This invention describes a novel method for the production of a CMP
CC (cytidine monophosphate)-sialate synthetase protein which comprises
CC culturing a prokaryotic host organism transformed with an expression
CC vector containing a structural gene for the protein, where the start
CC codon of the structural gene is 8-12 bases downstream from a ribosome
CC binding site (RBS). The protein is useful for producing CMP-activated N-
CC acetylneuraminic acid analogs by reacting the analogs with cytidine
CC triphosphate in the presence of the protein and for producing non-
CC naturally sialylated oligosaccharides and glycoconjugates, which are of

CC therapeutic interest because of their sialidase resistance. The protein
 CC is also useful for detecting sialyl acceptors by reacting the acceptor
 CC with a fluorescent sialic acid derivative in the presence of the protein.
 CC cytidine triphosphate and a sialyl transferase and detecting the
 CC fluorescently modified acceptor by selective chromatography. This
 CC sequence encodes a CMP (cytidine monophosphate)-sialate synthetase
 CC protein, sialB, isolated from *Neisseria meningitidis* serogroup B
 CC
 SQ Sequence 687 BP; 218 A; 154 C; 122 G; 193 T; 0 U; 0 Other;
 Query Match 60.0%; Score 27; DB 3; Length 687;
 Best Local Similarity 100.0%; Pred. No. 0.5;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 19 GCTTTCCTTGATTAAGATGTTTC 45
 DB 684 GCTTTCCTTGATTAAGATGTTTC 658
 RESULT 4
 ID AAA81502 standard; DNA; 65632 BP.
 AC AAA81502;
 XX
 DT 04-DEC-2000 (first entry)
 DE N. meningitidis partial DNA sequence gnm_49 SEQ ID NO:49.
 XX
 KM *Neisseria meningitidis*; *Neisseria gonorrhoeae*; genome; immunogenic;
 KM antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KM *Meningococcus B*; MenB; ds.
 XX
 OS *Neisseria meningitidis*.
 XX
 PN WO200022430-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99WO-US023573.
 XX
 PR 09-OCT-1998; 98US-0103794P.
 PR 30-APR-1999; 99US-0132068P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
 PI Masigmt V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappolli R, Pizzi M;
 XX
 DR WPI; 2000-318079/27.
 XX
 PT Isolated nucleotide sequences of *Neisseria meningitidis* which can be used
 PT in the diagnosis and treatment of *N. meningitidis* infection and other
 PT *Neisseria* infections, for example, *N. gonorrhoea*.
 XX
 PS Claim 7, Page 1331-1350; 1760pp; English.
 XX
 CC The present invention describes methods of obtaining immunogenic proteins
 CC from *Neisseria* genomic sequences. AAA81453 to AAA82414 represent
 CC specifically claimed *Neisseria meningitidis* genomic DNA sequences;
 CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent *Neisseria* DNA
 CC sequences and their corresponding proteins; AAA81254 to AAA81259 and
 CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
 CC *Neisseria meningitidis* DNA sequences; and AAA81322 to AAA81452 represent
 CC *Neisseria meningitidis* Mem polynucleotide ORF sequences, which are all
 CC used in the exemplification of the present invention. The nucleic acid
 CC sequences, protein sequences, and antibodies against them, can be used in
 CC the manufacture of a composition. The composition can be used as a
 CC medicament (or in the manufacture of a medicament) for treating,
 CC preventing or diagnosing infection due to *Neisseria* bacteria. For
 CC example, some of the identified proteins could be components of vaccines
 CC against *Meningococcus B*; against all serotypes; and/or against all

CC pathogenic *Neisseriae*, identification of sequences from the bacterium
 CC will also facilitate production of biological probes, particularly
 CC organism-specific probes. Attempts to make efficacious *Meningococcus B*
 CC vaccines have failed mainly due to antigen tolerance. Multivalent
 CC vaccines have also been tried but none have successfully overcome
 CC antigenic variability. The provision of further, complete sequences may
 CC provide an opportunity to identify secreted or surface exposed proteins
 CC that may be presumed targets for the immune system and which are not
 CC antigenically variable or at least more conserved than other more
 CC variable regions
 CC
 SQ Sequence 65632 BP; 16704 A; 16327 C; 15474 G; 17126 T; 0 U; 1 Other;
 Query Match 60.0%; Score 27; DB 3; Length 65632;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 19 GCTTTCCTTGATTAAGATGTTTC 45
 DB 55738 GCTTTCCTTGATTAAGATGTTTC 55712
 RESULT 5
 AAA81490_00
 WP Sequence split into 15 fragments LOCUS AAA81490 Accession Aaa81490
 WP Fragment Name Begin End
 WP AAA81490_00 1 110000
 WP AAA81490_01 100001 210000
 WP AAA81490_02 200001 310000
 WP AAA81490_03 300001 410000
 WP AAA81490_04 400001 510000
 WP AAA81490_05 500001 610000
 WP AAA81490_06 600001 710000
 WP AAA81490_07 700001 810000
 WP AAA81490_08 800001 910000
 WP AAA81490_09 900001 1010000
 WP AAA81490_10 1000001 1110000
 WP AAA81490_11 1100001 1210000
 WP AAA81490_12 1200001 1310000
 WP AAA81490_13 1300001 1410000
 WP AAA81490_14 1400001 1437668
 ID AAA81490 standard; DNA; 1437668 BP.
 XX
 AC AAA81490;
 XX
 DT 04-DEC-2000 (first entry)
 DE N. meningitidis B full length genome DNA sequence SEQ ID NO:1068.
 XX
 KM *Neisseria meningitidis*; *Neisseria gonorrhoeae*; genome; immunogenic;
 KM antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KM *Meningococcus B*; MenB; ds.
 XX
 OS *Neisseria meningitidis*.
 XX
 PN WO200022430-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99WO-US023573.
 XX
 PR 09-OCT-1998; 98US-0103794P.
 PR 30-APR-1999; 99US-0132068P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
 PI Masigmt V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappolli R, Pizzi M;
 XX
 DR WPI; 2000-318079/27.
 XX
 PT Isolated nucleotide sequences of *Neisseria meningitidis* which can be used

PT in the diagnosis and treatment of *N. meningitidis* infection and other
PT Neisserial infections, for example, *N.gonorrhoea*.

XX Claim 7; Page 866-1272; 1760pp; English.

CC The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AAB1453 to AAB2414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAB1260 to AAB81303 and AAB25620 to AAB25663 represent Neisseria DNA
CC sequences and their corresponding proteins; AAB1254 to AAB81259 and
CC AAB1304 to AAB81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAB81322 to AAB1452 represent
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against *Meningococcus B*; against all serotypes; and/or against all
CC pathogenic Neisseriae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious *Meningococcus B*
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions

XX SQ Sequence 1437668 BP; 344338A; 353206C; 385074G; 355045T; 0U; 50ther;

Query Match 60.0%; Score 27; DB 3; Length 110000;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 GCTTTCCTTGATTAAGAATGTTTC 45

DB 78420 GCTTTCCTTGATTAAGAATGTTTC 78446

RESULT 6
AAB21544
ID AAB21544 standard; DNA; 349980 BP.

XX AC AAB21544;

XX DT 13-MAR-2001 (first entry)

XX DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:1.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
XX ds.

OS Neisseria meningitidis.

XX PN WO200066791-A1.

XX PD 09-NOV-2000.

XX PF 08-MAR-2000; 2000MO-US005928.

XX PR 30-APR-1999; 199US-0132068P.

XX PR 08-OCT-1999; 99MO-US023573.

XX PR 28-FEB-2000; 2000GB-00004695.

XX PA (CHIR) CHIRON CORP.

XX PA (GENO-) INST GENOMIC RES.

XX PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scariselli M, Scarlato V;

PI Rappuoli R, Frazer CM, Grandi G;
XX WPI; 2000-647603/62.

XX PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections.

XX PS Claim 7; Appendix A; 692pp; English.

XX The present invention describes the full length genome of Neisseria
CC meningitidis B (NMB). The sequences in AAB21544 and AAB21607 to AAB21613
CC represent fragments of the NMB genomic sequence, as the sequence was too
CC long to go in a record on its own it was split into 8 sequences which
CC overlap each other at the beginning and end of each sequence by 49980 bp
CC (i.e. the last 49980 bp of AAB21544 is repeated at the beginning of
CC AAB21607, the last 49980 bp of AAB21607 are repeated at the beginning of
CC AAB21608, and so on). AAB21545 to AAB21588 encode the Neisseria proteins
CC given in AAB58550 to AAB58593, and AAB21589 to AAB21606 represent PCR
CC primers which are used in the exemplification of the present invention.
CC The NMB genome and fragments from it have antibacterial activity, and can
CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
CC and/or antibodies which binds to the proteins can be used in compositions
CC for treating or preventing infection due to Neisserial bacteria or as a
CC diagnostic reagent for detecting the presence of Neisserial bacteria or
CC of antibodies raised to Neisserial bacteria. Computers, computer memory,
CC computer storage medium or computer databases can be used in a search to
CC identify open reading frames (ORFs) or coding sequences within the NMB
CC genome. The DNA sequences provide further opportunities to find antigenic
CC or immunogenic proteins which are more effective in vaccines than the
CC outer membrane proteins currently used

XX SQ Sequence 349980 BP; 83241 A; 85091 C; 95206 G; 86442 T; 0 U; 0 Other;

Query Match 60.0%; Score 27; DB 3; Length 349980;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 GCTTTCCTTGATTAAGAATGTTTC 45

DB 78420 GCTTTCCTTGATTAAGAATGTTTC 78446

RESULT 7
ABV36932
ID ABV36932 standard; cDNA; 367 BP.

XX AC ABV36932;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 36923.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001MO-US005171.

XX PR 17-FEB-2000; 2000US-018319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JB;

XX MPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 7608; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 CC
 SO Sequence 367 BP; 88 A; 82 C; 98 G; 97 T; 0 U; 2 Other;

Query Match 53.3%; Score 24; DB 5; Length 367;
 Best Local Similarity 75.0%; Pred. No. 8.1;
 Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAGAAATTCGGCCACCGCTTCTTGATTAAGAATGTTT 43
 |||||
 DB 271 CACAATTAGCCTCTACTCTTCTTGATTAAGAATGTTT 310

RESULT 8
 ABV42556
 ID ABV42556 standard; cDNA; 367 BP.
 XX
 AC ABV42556;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 42547.
 XX
 KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KM pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2;
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MILL-) MILLENNIUM PRDICTIVE MEDICINE INC.
 PI Schlegel R, Endege WO, Monahan JR;
 XX
 DR MPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 8517; 11750pp; English.
 XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 CC
 SO Sequence 367 BP; 88 A; 82 C; 98 G; 97 T; 0 U; 2 Other;

Query Match 53.3%; Score 24; DB 5; Length 367;
 Best Local Similarity 75.0%; Pred. No. 8.1;
 Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CAGAAATTCGGCCACCGCTTCTTGATTAAGAATGTTT 43
 |||||
 DB 271 CACAATTAGCCTCTACTCTTCTTGATTAAGAATGTTT 310

RESULT 9
 ADT77703
 ID ADT77703 standard; DNA; 31 BP.
 XX
 AC ADT77703;
 XX
 DT 13-JAN-2005 (first entry)
 XX
 DB Neisseria meningitidis capsule synx gene PCR primer capRI.
 XX
 KM Vaccine; attenuation; synx gene; antibacterial; PCR; primer; ss.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO2004089408-A2.
 XX
 PD 21-OCT-2004.
 XX
 PF 07-APR-2004; 2004WO-GB001504.
 XX
 PR 07-APR-2003; 2003GB-00008691.
 XX
 PA (XENO-) XENOVA RES LTD.
 PI McLean CS, Keen SW, Martin GM;
 XX
 DR MPI; 2004-748603/73.
 XX
 PT New attenuated mutant Neisseria bacterium comprising an auxotrophic
 PT attenuating mutation, a capsule mutation and a mutation which reduces
 PT bacterial recombination or exogenous DNA uptake, used as vaccine against
 PT Neisseria infection.
 XX
 PS Example; SEQ ID NO 9; 40pp; English.
 XX
 CC The present sequence is that of PCR primer capRI for the Neisseria
 CC meningitidis synx gene, which is involved in bacterial capsule formation.
 CC PCR amplification of strain B16B6 genomic DNA using capRI ADT77702 and
 CC capRI generated a 3.1 kb fragment extending from the 5' end of the slab
 CC gene to midway through the ctrA gene. This PCR product was used as
 CC template for the PCR amplification of left and right hand flanking
 CC regions of the synx gene. Primer capRI was used with primer capdelFI
 CC ADT77705 to generate a 668 bp right hand flanking region. This was used
 CC in the generation of plasmids for synx deletion and hence in the
 CC construction of an arCB/synx/reca triple mutant live attenuated vaccine
 CC strain. The invention provides an attenuated strain of a Neisseria
 CC bacterium, including N. meningitidis, that retains immunogenicity and
 CC ability to grow in culture, without causing neisserial disease in
 CC subjects to which it is administered as a vaccine. The attenuated

CC Neisseria bacterium, or a live or killed preparation of it, has the
 CC following mutations: (a) an auxotrophic attenuation mutation, e.g. an
 CC ArcA or an ArcB mutation; (b) a capsule mutation that affects capsule
 CC integrity and/or causes the capsule to be of reduced thickness or absent,
 CC e.g. a mutation in the *synX* gene or the *gale* gene; and (c) a mutation
 CC which reduces bacterial recombination or exogenous DNA uptake, e.g. a
 CC mutation in a bacterial recombinase gene such as *recA*, or alternatively a
 CC capsid mutation. A claimed pharmaceutical composition for use as a vaccine
 CC comprises the live or killed mutant *Neisseria* bacterium.

XX Sequence 31 BP; 8 A; 5 C; 6 G; 12 T; 0 U; 0 Other;

Query Match 52.4%; Score 23.6; DB 13; Length 31;

Best Local Similarity 86.7%; Pred. No. 6.7;
 Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 10 TCCGACACGCGCTTCTTGATTAAGATGT 39
 Db 2 TCCGATATCGCTTCTTGATTAAGAT 31

RESULT 10

ABL80063/C
 ID ABL80063 standard; cDNA; 351 BP.

XX ABL80063;

XX 17-MAY-2002 (first entry)

DE Human ovarian cancer related cDNA clone SEQ ID NO:3041.

KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX MO200192581-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US017756.

XX 26-MAY-2000; 2000US-0207484P.

XX (CORI-) CORIXA CORP.

PI Algate PA, Harlocker SL, Jones R;

DR MPI; 2002-122075/16.

PT Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PT polypeptide.

XX Claim 1; SEQ ID NO 3041; 489pp; English.

XX The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) having a cDNA sequence (SI)
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (SI)
 CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
 CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
 CC or antigen presenting cells that express (II) (I) has cytostatic
 CC activity. An oligonucleotide (IV) that hybridises to (SI) can be used for
 CC detecting ovarian cancer in a patient's biological sample preferably
 CC serum or ovarian tissue. The method comprises contacting a polynucleotide
 CC sample from a patient with (IV), detecting the amount of polynucleotide
 CC hybridising to (IV) and comparing the amount to a predetermined cutoff
 CC value and thereby detecting ovarian cancer in the patient, where the
 CC amount of polynucleotide hybridising to (IV) is detected preferably by
 CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
 CC useful for stimulating and/or expanding T cells specific for an ovarian
 CC tumor protein comprising contacting T cells with (III) or (II). (III) is
 CC useful in design and preparation of ribozyme molecules for inhibiting

CC expression of the tumour polypeptides and proteins in tumour cells; and
 CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
 CC library using well known techniques

XX Sequence 351 BP; 74 A; 106 C; 91 G; 80 T; 0 U; 0 Other;

Query Match 52.0%; Score 23.4; DB 6; Length 351;

Best Local Similarity 73.2%; Pred. No. 14;
 Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCGCACGCGCTTCTTGATTAAGATGT 42
 Db 240 GACAGATTCCGCGCGCTTCTTGATTAAGATGT 200

RESULT 11

ADR07328/C
 ID ADR07328 standard; cDNA; 3812 BP.

XX ADR07328;

XX 04-NOV-2004 (first entry)

DE Full length human cDNA useful for treating neurological disease Seq 834.

KW gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
 KW osteoporosis; neurological disease; Alzheimer's disease;
 KW Parkinson's disease; dementia; short memory; cancer;
 KW sense or motor function; emotional reaction; fear response; panic;
 KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
 KW tranquilliser.

XX Homo sapiens.

XX EPI447413-A2.

XX 18-AUG-2004.

XX 12-FEB-2004; 2004EP-00003145.

XX 14-FEB-2003; 2003JP-00102207.

XX 09-MAY-2003; 2003JP-00131452.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;

PI Wakamatsu A, Ishii S, Nagai K, Irie R;

DR MPI; 2004-583265/57.

DR P-PDB; ADR09284.

PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 834; 2686pp; English.

XX This invention relates to novel, isolated full length human cDNA
 CC molecules and the encoded proteins thereof. Specifically, it refers to
 CC cDNA clones obtained by an oligo-capping method, where none of these
 CC clones are identical to any known human mRNAs. The present invention
 CC describes an immunoassay to identify agonists and antagonists, as well as
 CC antibodies, antisense molecules and siRNAs that can all be used to bind
 CC to and modulate expression of the cDNA molecules. As such, these
 CC molecules are useful for diagnostic markers or therapeutic targets for
 CC the various diseases or morbid states. In particular, they are useful in
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
 CC disease, Parkinson's disease, dementia, short memory and various cancers,
 CC as well as for maintaining equilibrium of sense or motor function, and
 CC for treating emotional reaction, fear response and panic. Accordingly,
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
 CC cyostatic and tranquilliser activities. This polynucleotide is a full
 CC length human cDNA sequence of the invention. NOTE: This sequence is not
 CC given in the sequence listing of the specification but can be obtained on

CC CD-ROM from the European Patent Office, Vienna Sub-office.

XX Sequence 3812 bp; 778 A; 1002 C; 1108 G; 924 T; 0 U; 0 Other;

XX SQ

Query Match

Best Local Similarity 52.0%; Score 23.4; DB 13; Length 3812;

Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATGTT 42

DB 2857 GACACGATTCCGCTCCACTGTCTCTCAGTAACAGCTT 2817

RESULT 12

ADQ63001/c

ID

ADQ63001 standard; cDNA; 4470 BP.

XX ADQ63001;

XX 07-OCT-2004 (first entry)

XX Novel human cDNA sequence #162.

XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
XX cytosolic; gene therapy; diagnostic marker; morbid state; osteoporosis;
XX neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
XX cancer.

XX Homo sapiens.

XX EP1440981-A2.

XX 28-JUL-2004.

XX 21-JAN-2004; 2004EP-00001196.

XX 21-JAN-2003; 2003JP-00102206.

XX 09-MAY-2003; 2003JP-00131392.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

XX Yamamoto J, Isono Y, Nagai K, Irie R;

XX MPI; 2004-535376/52.

XX P-PSDB; ADQ65189.

XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 162; 2449bp; English.

XX The invention relates to 2495 novel polynucleotides (1) and their encoded
XX polypeptides, sequences hybridizing to these nucleotides, sequences
XX encoding partial polypeptides and sequences having 70% or 90% identity to
XX the nucleotide and protein sequences. The nucleotides and polypeptides
XX are useful as diagnostic markers or therapeutic target for the diseases
XX or morbid states. They are also useful for treating osteoporosis,
XX neurological diseases, Alzheimer's diseases, Parkinson's diseases,
XX dementia and various cancers. This sequence corresponds to a nucleotide
XX sequence of the invention.

XX Sequence 4470 BP; 909 A; 1191 C; 1285 G; 1085 T; 0 U; 0 Other;

XX Query Match

Best Local Similarity 52.0%; Score 23.4; DB 12; Length 4470;

Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATGTT 42

DB 3512 GACAGATTCCGCTCCACTGTCTCTCAGTAACAGCTT 3472

RESULT 13

ADH85460

ID

ADH85460 standard; DNA; 720 BP.

XX ADH85460;

XX 22-APR-2004 (first entry)

XX Enterococcus faecalis polynucleotide #3345.

XX Enterococcus faecalis infection; transcription regulatory element;
XX antibacterial; gene; ds.

XX Enterococcus faecalis.

XX US6617156-B1.

XX 09-SEP-2003.

XX 13-AUG-1998; 98US-00134000.

XX 15-AUG-1997; 97US-0055778P.

XX (DOUC/) DOUCETTE-STAMM L A.

XX (BUSH/) BUSH D.

XX Doucette-Stamm LA, Bush D;

XX MPI; 2003-895394/82.

XX P-PSDB; ADH88865.

XX New nucleic acid comprising a sequence encoding an Enterococcus faecalis
XX polypeptide, useful for preparing a composition for diagnosing or
XX treating E. faecalis infection.

XX Disclosure; SEQ ID NO 3345; 193bp; English.

XX The invention relates to Enterococcus faecalis polynucleotides and
XX polypeptides. The invention also relates to a recombinant expression
XX vector comprising a polynucleotide operably linked to a transcription
XX regulatory element, a cell comprising a recombinant vector, a method for
XX producing an E. faecalis polypeptide, an isolated nucleic acid comprising
XX a sequence not given in the specification, a recombinant vector
XX comprising the nucleic acid and a cell comprising the recombinant vector.
XX The polynucleotides can be used to detect the presence of E. faecalis in
XX a sample. The sequences are useful for preparing a composition for
XX diagnosing or treating Enterococcus faecalis infection. This sequence
XX represents an E. faecalis polynucleotide of the invention.

XX Sequence 720 BP; 244 A; 93 C; 147 G; 236 T; 0 U; 0 Other;

XX Query Match

Best Local Similarity 51.6%; Score 23.2; DB 10; Length 720;

XX Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGA 37

DB 36 GAAAGATTTCGATTCCTTCTTGATTAAGA 71

RESULT 14

AAZ24471

ID

AAZ24471 standard; DNA; 35 BP.

XX AAZ24471;

XX 17-FEB-2000 (first entry)

XX N. meningitidis siab PCR primer 2.

XX CMP sialylate synthetase; siab; cytidine monophosphate; detection;
XX CMP-activated N-acetylneuraminic acid; sialyl acceptor; PCR primer; ss.

OS Synthetic.
 OS Neisseria meningitidis.
 XX
 XX DEL913206-A1.
 XX
 XX 07-OCT-1999.
 PD
 XX
 PF 24-MAR-1999; 99DE-01013206.
 XX
 PR 26-MAR-1998; 98DE-01013426.
 XX
 PA (FESS/) FESSNER W D.
 XX
 PI Fessner W, Knöfret M;
 XX
 DR WPI; 2000-024923/03.
 XX
 XX Production of CMP-sialate synthetase protein useful for producing CMP-
 PT activated N-acetylneuraminic acid analogs and non-naturally sialylated
 PR oligosaccharides and glycoconjugates.
 XX
 XX
 PS Claim 5; Page 18; 26pp; German.
 XX
 CC This invention describes a novel method for the production of a CMP
 CC (cytidine monophosphate)-sialate synthetase protein which comprises
 CC culturing a prokaryotic host organism transformed with an expression
 CC vector containing a structural gene for the protein, where the start
 CC codon of the structural gene is 8-12 bases downstream from a ribosome
 CC binding site (RBS). The protein is useful for producing CMP-activated N-
 CC acetylneuraminic acid analogs by reacting the analogs with cytidine
 CC triphosphate in the presence of the protein and for producing non-
 CC naturally sialylated oligosaccharides and glycoconjugates, which are of
 CC therapeutic interest because of their sialidase resistance. The protein
 CC is also useful for detecting sialyl acceptors by reacting the acceptor
 CC with a fluorescent sialic acid derivative in the presence of the protein,
 CC cytidine triphosphate and a sialyl transferase and detecting the
 CC fluorescence modified acceptor by selective chromatography. This
 CC sequence represents a PCR primer used in the amplification of a CMP
 CC (cytidine monophosphate)-sialate synthetase protein, s1ab, isolated from
 CC Neisseria meningitidis serogroup B
 XX
 SQ Sequence 35 BP; 7 A; 6 C; 8 G; 14 T; 0 U; 0 Other;
 XX
 Query Match 50.7%; Score 22.8; DB 3; Length 35;
 Best Local Similarity 79.4%; Pred. No. 15;
 Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 7 AATTCGCGACCGCTTCTTGATTAAGAATG 40
 DB 2 ATTCGCGGTTAGCTTCTTGATTAAGAATG 35
 XX
 RESULT 15
 ADM02496
 ID ADM02496 standard; cDNA; 3371 BP.
 XX
 AC ADM02496;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Human cDNA of the invention SEQ ID NO:1181.
 XX
 KW ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
 XX
 OS Homo sapiens.
 XX
 XX EPI347046-A1.
 XX
 PD 24-SEP-2003.
 XX
 PF 12-APR-2002; 2002EP-00008400.
 XX
 PR 22-MAR-2002; 2002JP-00137785.

XX
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuho Y;
 XX
 DR WPI; 2003-723558/69.
 DR P-PSDB; ADM04939.
 XX
 XX New polynucleotides and polypeptides are useful in gene therapy, for
 PT developing a diagnostic marker or medicines for regulating their
 PT expression and activity, or as a target of gene therapy.
 XX
 PS Claim 1; SEQ ID NO 1181; 305pp; English.
 XX
 CC The invention relates to a novel human polynucleotide and the encoded
 CC polypeptide. A polynucleotide of the invention ADM06202-ADM06773 is useful
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
 CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a
 CC cDNA sequence of the invention.
 XX
 SQ Sequence 3371 BP; 833 A; 892 C; 880 G; 766 T; 0 U; 0 Other;
 XX
 Query Match 50.7%; Score 22.8; DB 11; Length 3371;
 Best Local Similarity 71.4%; Pred. No. 43;
 Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 3 ACAGAAATTCGCGACCGCTTCTTGATTAAGAATGTTT 44
 DB 1378 ACAGAAATTCGCGACCGCTTCTTGATTAAGAATGTTT 1419
 XX
 Search completed: April 7, 2006, 16:17:21
 Job time : 55.7604 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:00:12 ; Search time 353.403 Seconds
(without alignments)
5957.564 Million cell updates/sec

Title: US-09-211-691-4

Perfect score: 45
Sequence: 1 cgacagatccgcacacgc.....ctgtgattaagaatgtttc 45

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4108325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gse1:*
10: gb_gse2:*
11: gb_gse3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24.6	54.7	672	9	AZ485731 1M0313018
2	24.6	54.7	1958	4	AK017055 Mus muscu
3	24.4	54.2	625	9	BH333088 BH333088
4	24.4	54.2	833	2	BE641022 Cr12.2 A2
5	24.4	54.2	1022	5	BUI58055 AGENCOURT
6	23.8	52.9	283	10	CE688260 tigr-gse-
7	23.8	52.9	500	3	BP188097 BP188097
8	23.8	52.9	815	5	BK698840 BK698840
9	23.6	52.4	587	9	BH333091 CH230-59K
10	23.6	52.4	780	9	CC929966 ZMMB053
11	23.6	52.4	888	4	CNS09115 BX063945 Single re
12	23.6	52.4	962	10	CL994393 ZMMB000
13	23.6	52.4	1087	10	CW001860 ZMMB000
14	23.4	52.0	284	1	BB159100 BB159100
15	23.4	52.0	351	1	AA292589 ZC31F01.x
16	23.4	52.0	416	3	BM694767 UI-E-C11-
17	23.4	52.0	433	7	CO997065 pam01-15m
18	23.4	52.0	501	10	CG603013 OST277031
19	23.4	52.0	540	8	DN648303 G7023.62
20	23.4	52.0	576	7	CV002745 pam01-19m
21	23.4	52.0	592	7	CK534070 rswg0.00
22	23.4	52.0	634	10	AG095231 Pan tlogl

23	23.4	52.0	653	8	CX185987	CX185987 B01.45-76
24	23.4	52.0	670	6	CA924883	CA924883 MTU7T.P1
25	23.4	52.0	674	2	BC911192	BC911192 602807115
26	23.4	52.0	754	4	DN054816	DN054816 UGI_CABAB
27	23.4	52.0	2035	4	BSM807446	BX647302 Homo sapi
28	23.2	51.6	407	5	BQ475193	BQ475193 ciciendela
29	23.2	51.6	430	1	AM659916	AM659916 97576 MAR
30	23.2	51.6	554	8	DR872146	DR872146 JGI_CABG1
31	23.2	51.6	566	8	B2760599	B2760599 623_012.B
32	23.2	51.6	741	9	BH498370	BH498370 BQJ030TF
33	23.2	51.6	750	10	CG852499	CG852499 ZMMB0034
34	23	51.1	260	1	AV360233	AV360233 AV360233
35	23	51.1	393	9	AO168424	AO168424 HS_3151.B
36	23	51.1	533	6	CB422282	CB422282 595372 MA
37	23	51.1	539	6	CA968909	CA968909 CUL00a10
38	23	51.1	559	6	CB421928	CB421928 594988 MA
39	23	51.1	612	8	DT110051	DT110051 JGI_ANN06
40	23	51.1	612	8	DT115631	DT115631 JGI_ANN03
41	23	51.1	676	6	CB429698	CB429698 605532 MA
42	23	51.1	679	8	DN543254	DN543254 1392090 M
43	23	51.1	680	8	DN541153	DN541153 1386690 M
44	23	51.1	729	7	CK247530	CK247530 EST731167
45	23	51.1	729	11	CR312301	CR312301 mte1-34P2

ALIGNMENTS

RESULT 1
AZ485731
LOCUS
DEFINITION
1M0313018F Mouse 10kb plasmid UGCGM library Mus musculus genomic
clone UGCGM0313018 F, genomic survey sequence.

ACCESSION
AZ485731
VERSION
AZ485731.1 GI:10651816
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)

REFERENCE

AUTHORS

Dunn,D., Moyagi,A., Barber,M., Beacorn,T., Duval,B., Haml,C.,
Islam,H., Longacre,S., Mahmoud,E., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingley,A., von
Niederhausern,A. and Wright,D.,Weiss,R.
1 (bases 1 to 672)
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0313 row: 0 column: 18
Seq primer: CGTTGTAAACGACGCGCCAGT
Class: plasmid ends
High quality sequence strop: 672.
Location/Qualifiers
1..672
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="UGCGM0313018"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/clone_lib="Mouse 10kb plasmid UGCGM library"
/note="Vector: PMD42nv, Purified genomic DNA from M."

FEATURES

source

polyA_signal 1943.1948
/note="putative"
polyA_site 1958
/note="putative"

Query Match 54.7%; Score 24.6; DB 4; Length 1958;
Best Local Similarity 76.9%; Pred. No. 1.3e+02;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 6 GAATTCGGCCACCGCTTCTCTGTGATTGAAGATGTTT 44
| | | | | | | | | | | | | | | | | | | | | |
Db 849 GGATTCGCCCAATGCTTTCATGTGAACCTGAATGTTT 887

RESULT 3
LOCUS BH333088 625 bp DNA linear GSS 03-DEC-2001
DEFINITION CH230-59K11.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
ACCESSION BH333088
VERSION BH333088.1 GI:117263802
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 625)
Zhao, S., Shetty, J., Shateman, S., Tsagaye, G., Geer, K.,
Shvartsbeyn, A., Gebregiorgis, E., Overton, L., Russell, D., Chen, D.,
Riggs, F., de Jong, P. and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 EcORI segment
Unpublished (1999)

TITLE JOURNAL
AUTHORS Other-GSSs: CH230-59K11.TV
COMMENT Department of Bukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdjong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orering/information.htm). BAC end
pages: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 59 row: K column: 11
Seq primer: SP6
Class: BAC ends.

FEATURES
source Location/Qualifiers
1..625

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-59K11"
/sex="Female"
/cell_type="Brain"
/clone_1lb="CHORI-230 Segment 1"
/note="Vector: PTARBA2.1; Site 1: EcORI; Site 2: EcORI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN

Query Match 54.2%; Score 24.4; DB 9; Length 625;
Best Local Similarity 73.8%; Pred. No. 1.3e+02;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 3 ACAGATTCCGCGCCGCTTCTCTGTGATTGAAGATGTTT 44
| | | | | | | | | | | | | | | | | | | | | |
Db 39 ACAGATACCGCTACCTTACCTTGTATCAACAGTACTT 80

RESULT 4

BE641022/c 833 bp mRNA linear EST 01-SEP-2000
LOCUS Cr12.2 A24 SP6 Ceratopteris Spore Library Ceratopteris richardii
DEFINITION cDNA clone Cr12.2_A24 5', mRNA sequence.

ACCESSION BE641022
VERSION BE641022.1 GI:9958681
KEYWORDS EST.
SOURCE Ceratopteris richardii
ORGANISM Ceratopteris richardii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Moniliformes; Filicophyta; Filicopsida; Filicales; Pteridaceae;
Ceratopteris.

REFERENCE 1 (bases 1 to 833)
Chatterjee, A., San Miguel, P., Stout, S.C., Banks, J. and Roux, S.J.
Expressed sequence tags of cDNA clones from a C. richardii library
Unpublished (2000)
COMMENT Contact: Roux SJ
Section of Molecular Cell and Developmental Biology
University of Texas
Biology Building, Room 16, Austin, TX 78712, USA
Tel: 512 471 4238
Fax: 512 232 3402
Email: sroux@utcc.utexas.edu
Plate: Cr12.2 row: A column: 24
Seq primer: SP6.

FEATURES
source Location/Qualifiers
1..833

/organism="Ceratopteris richardii"
/mol_type="mRNA"
/cultivar="Brogan"
/db_xref="taxon:49495"
/clone="Cr12.2 A24"
/tissue_type="Gametophyte"
/cell_type="Spore"
/dev_stage="20 hours after germination initiation"
/clone_1lb="Ceratopteris Spore Library"
/note="Vector: pCMVSPORT6; EST sequence from cDNA library.
cDNA library constructed from mRNA isolated from C.
richardii spores that had developed for 20 hours after
their germination had been initiated by white light."

ORIGIN

Query Match 54.2%; Score 24.4; DB 2; Length 833;
Best Local Similarity 73.8%; Pred. No. 1.3e+02;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GACGAAATTCGCGCCGCTTCTCTGTGATTGAAGATGTTT 43
| | | | | | | | | | | | | | | | | | | | | |
Db 513 GCCAAATTCGCCCTCTCTCTCTCTCTGCGCATTAAGATTTT 472

RESULT 5
LOCUS BU158055 1022 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT 8043170 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6084690
5', mRNA sequence.
ACCESSION BU158055
VERSION BU158055.1 GI:22671965
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE 1 (bases 1 to 1022)
NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

REFERENCE
1 (bases 1 to 815)
AUTHORS
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE
Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL
Unpublished (2003)
COMMENT
Contact: Croning MDR

Xenopodinae; Xenopus; Silurana.
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu085603.q1k77
Sequencing primer: T7
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1..815
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TNeu085603"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."

FEATURES

source

ORIGIN

Query Match 52.9%; Score 23.8; DB 5; Length 815;
Best Local Similarity 72.1%; Pred. No. 2.3e+02;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 3 ACAGATTCCGACCGCTTCTTGATTAAGATGTTTC 45
Db 312 AAAAATTTCAGACACTTATCTTGCTGTAATAATCTTTC 354

RESULT 9
BH333091 587 bp DNA linear GSS 03-DEC-2001
LOCUS CH230-59K12.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
DEFINITION BH333091
ACCESSION BH333091
VERSION BH333091.1 GI:17263805
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
1 (bases 1 to 587)
Zhao,S., Shetty,J., Shateman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregeorgis,B., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other GSSs: CH230-59K12.TV
Contact: Shaying Zhao
Department of Bukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tcd/bac_ends/rat/bac_end_intro.html
Plate: 59 row: K column: 12
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..587
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SmHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-59K12"
/sex="Female"
/cell_type="Brain"
/clone_1lb="CHORI-230 Segment 1"
/note="Vector: pTARBA2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SmHsd/MCW) BAC library produced by
Pieter de Jong"

FEATURES

source

ORIGIN

Query Match 52.4%; Score 23.6; DB 9; Length 587;
Best Local Similarity 76.3%; Pred. No. 2.6e+02;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3 ACAGATTCCGACCGCTTCTTGATTAAGATG 40
Db 39 ACAGATACCCTTACTTACCTTGATTAACAGTG 76

RESULT 10
CC929966 780 bp DNA linear GSS 11-AUG-2003
LOCUS ZMMBc0538U06r ZMMBc Zea mays genomic clone ZMMBc0538U06 3'
DEFINITION CC929966
ACCESSION CC929966
VERSION CC929966.1 GI:33581064
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 780)
Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Rouzaud,K., Fuks,G., Yu,Y., Wang,R. and Messing,J.
Sequencing of the maize genome at PGIR (2003b)
Unpublished (2003)
Contact: Bharti,A.K.
Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Wakeman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 510.
Location/Qualifiers
1..780
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBc0538U06"
/lab_host="E. coli DH10B"
/clone_1lb="ZMMBc"
/note="Vector: pTARBA2.1; Site_1: BamHI; Site_2: BamHI"

FEATURES

source

ORIGIN

Query Match 52.4%; Score 23.6; DB 9; Length 780;
Best Local Similarity 76.3%; Pred. No. 2.7e+02;

Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 8 ATTCCGACCGCTTCTTGATTAAGATGTTTC 45
 |||||
 445 ATTGGCCATGGCTTCTTGATTAAGATGTTTC 482

RESULT 11
 CNS09L15/c 888 bp mRNA linear HTC 22-APR-2005
 LOCUS Single read from an extremity of a full-length cDNA made from
 DEFINITION Anopheles gambiae total adult females. 3-PRIME end of clone
 FK0AAC6AB06 of strain 6-9 of Anopheles gambiae (African malaria
 mosquito).

ACCESSION BX063945.1 GI:27637226
 VERSION HTc.
 KEYWORDS Anopheles gambiae (African malaria mosquito)
 SOURCE Anopheles gambiae
 ORGANISM Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 Culicidae; Anophelinae; Anopheles.

REFERENCE 1
 AUTHORS Gomez,S.M., Eiglmeier,K., Segurens,B., Dehoux,P., Couloux,A.,
 Scarpelli,C., Wincker,P., Weissbach,J., Brey,P.T. and Roth,C.W.
 TITLE Pilot Anopheles gambiae full-length cDNA study: sequencing and
 JOURNAL initial characterization of 35,575 clones
 PUBMED Genome Biol. 6 (4), R39 (2005)
 15833126
 2 (bases 1 to 888)
 REFERENCE Genoscope.
 AUTHORS Direct Submission
 TITLE Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 FEATURES
 source Location/Qualifiers
 1. .888
 /organism="Anopheles gambiae"
 /mol_type="mRNA"
 /strain="6-9"
 /db_xref="taxon:7165"
 /clone="FK0AAC6AB06"
 /plasmid="pME18S-FL"
 /note="end : 3-PRIME"

ORIGIN

Query Match 52.4%; Score 23.6; DB 4; Length 888;
 Best Local Similarity 76.3%; Pred. No. 2.8e+02;
 Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 3 ACAGAAATCCGACCGCTTCTTGATTAAGATG 40
 |||||
 581 ACCGAGCGCGGACCGCTTCTTGATTAAGATG 544

RESULT 12
 CL994393 962 bp DNA linear GSS 23-SEP-2004
 LOCUS ZMMBHf0006h02.r ZMMBHf Zea mays genomic clone ZMMBHf0006h02 3',
 DEFINITION genomic survey sequence.
 ACCESSION CL994393
 VERSION CL994393.1 GI:52562471
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 962)
 Ma,J., SanMiguel,P., Liu,R., Haller,K., Soderlund,C. and
 Bennetzen,J.
 TITLE ZMMBH sequences
 JOURNAL Unpublished (2004)

COMMENT Contact: Jeff Bennetzen
 Bennetzen Lab
 The University of Georgia
 Department of Genetics, C426a Life Sciences Building, Athens, GA
 30602, USA
 Tel: 706-542-3698
 Fax: 706-583-0972
 Email: maize@uga.edu
 Plate: 0006 row: h column: 02
 Class: BAC ends.
 FEATURES
 source Location/Qualifiers
 1. .962
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBHf0006h02"
 /tissue_type="immature ear"
 /dev_stage="6-8 weeks"
 /lab_host="DH10B"
 /clone_lib="ZMMBHf"
 /note="Vector: TOPOpcr4; Site_1: EcoRI; Site_2: EcoRI"

ORIGIN

Query Match 52.4%; Score 23.6; DB 10; Length 962;
 Best Local Similarity 76.3%; Pred. No. 2.8e+02;
 Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 3 ACAGAAATCCGACCGCTTCTTGATTAAGATG 40
 |||||
 913 ACAGAAATCCGCTTCTTGATTAAGATTAACAGG 950

RESULT 13
 CM001860 1087 bp DNA linear GSS 23-SEP-2004
 LOCUS ZMMBHg0002h24.f ZMMBHg Zea mays genomic clone ZMMBHg0002h24 5',
 DEFINITION genomic survey sequence.
 ACCESSION CM001860
 VERSION CM001860.1 GI:52576921
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 1087)
 Ma,J., SanMiguel,P., Liu,R., Haller,K., Soderlund,C. and
 Bennetzen,J.
 TITLE ZMMBH sequences
 JOURNAL Unpublished (2004)
 COMMENT Contact: Jeff Bennetzen
 Bennetzen Lab
 The University of Georgia
 Department of Genetics, C426a Life Sciences Building, Athens, GA
 30602, USA
 Tel: 706-542-3698
 Fax: 706-583-0972
 Email: maize@uga.edu
 Plate: 0002 row: h column: 24
 Class: BAC ends.
 FEATURES
 source Location/Qualifiers
 1. .1087
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBHg0002h24"
 /tissue_type="immature ear"
 /dev_stage="6-8 weeks"
 /lab_host="DH10B"
 /clone_lib="ZMMBHg"
 /note="Vector: TOPOpcr4; Site_1: EcoRI; Site_2: EcoRI"

ORIGIN

Query Match 52.4%; Score 23.6; DB 10; Length 1087;
 Best Local Similarity 76.3%; Pred. No. 2.9e+02;
 Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3 AGAATTCGCGCACCGCTTCTTGATTAAGAATG 40
 |||||
 Db 913 AGAATTCCTTCTTGCTTCTTGATTAACAAG 950
 *|||

RESULT 14 284 bp mRNA linear EST 29-JUN-2000
 BBI59100 RIKEN full-length enriched, 16 days neonate thymus Mus
 LOCUS BBI59100 musculus cDNA clone A130043N04 3', mRNA sequence.
 DEFINITION BBI59100
 ACCESSION BBI59100
 VERSION BBI59100.1 GI:8815030
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murinae; Mus.
 1 (bases 1 to 284)
 Komori, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,
 Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
 Hirokane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
 Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
 Kiyosawa, H., Kojima, Y., Kondo, S., Koyama, S., Kurahara, C.,
 Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,
 Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,
 Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Suganara, Y., Suzuki, H., Suzuki, H., Tagawa, A.,
 Takahashi, F., Tomimaga, N., Toyai, T., Tsunoda, Y., Watabiki, A.,
 Watanabe, S., Yamamura, T., Yamazaki, I., Yano, R., Yasunishi, A.,
 Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M., and
 Hayashizaki, Y.

TITLE RIKEN Mouse ESTs (Komori, H., et al.)
 JOURNAL Unpublished (2000)
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S.,
 Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 The most stable and thermally stable of thermolabile enzymes by
 trehalose and its application for the synthesis of full length
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh, M., Kitsuina, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
 Tomaru, Y., Carninci, P., Shibata, Y., Osawa, Y., Muramatsu, M.,
 Okazaki, Y., and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P., and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for
 further details.

FEATURES
 source Location/Qualifiers
 1..284
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="A130043N04"
 /tissue_type="thymus"
 /dev_stage="16 days neonate"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 16 days neonate"

ORIGIN

Query Match 52.0%; Score 23.4; DB 1; Length 284;
 Best Local Similarity 73.2%; Pred. No. 2.8e+02;
 Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 5 AGAATTCGCGCACCGCTTCTTGATTAAGAATGTTTC 45
 |||||
 Db 115 AGCTTCCCGCACGTTTCTTGATTAAGAATGTTTCC 155
 |||||

RESULT 15 351 bp mRNA linear EST 08-AUG-1997
 LOCUS AA292589
 DEFINITION z31f01.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
 IMAGE:723961 5', mRNA sequence.
 ACCESSION AA292589
 VERSION AA292589.1 GI:1940583
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 351)
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
 Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
 Hawkins, M., Hultman, M., Kucaba, T., Lacey, M., Le, M., Le, N.,
 Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
 Rohlfing, T., Scheinberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
 Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,
 and Marra, M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)
 8889549
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL, contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1162 Std Error: 0.00
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 340.

FEATURES
 source Location/Qualifiers
 1..351
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:723961"
 /sex="Female"
 /tissue_type="ovarian tumor"
 /lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares ovary tumor NBHOT"
 /note="Organ: ovary; Vector: pT73D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer (5'
 TGTTCACATCTGAGTGGAGCGCGCGCTTTTTTTTTTTTTTT 3').
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library constructed by Bento Soares and
 M.Fatima Bonaldo. "

ORIGIN

Query Match 52.0%; Score 23.4; DB 1; Length 351;
 Best Local Similarity 73.2%; Pred. No. 2.9e+02;
 Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGAATGT 42
 DB 240 GACAGATTCCGCCCTCCACTGTCCTTCAGTAACAGCTT 200

Search completed: April 7, 2006, 18:17:07
 Job time : 360.403 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 16:17:47 ; Search time 16.0996 Seconds
(without alignments)
4968.475 Million cell updates/sec

Title: US-09-211-691-4
Perfect score: 45.1
Sequence: 1 cgacagaattccgcacccgc.....ttgtgataagaatgtttcc 45

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/1-COMB.seq:*
2: /cgn2_6/prodata/1/ina/5-COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A-COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B-COMB.seq:*
5: /cgn2_6/prodata/1/ina/R-COMB.seq:*
6: /cgn2_6/prodata/1/ina/PC/US-COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP-COMB.seq:*
8: /cgn2_6/prodata/1/ina/RB-COMB.seq:*
9: /cgn2_6/prodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23.2	51.6	720	3	US-09-134-000C-3345
2	23.2	51.1	601	3	US-09-949-016-156329
3	23	51.1	101300	3	US-09-949-016-156329
4	22.6	50.2	364	3	US-09-513-999C-26046
5	22.4	49.8	2576	3	US-09-265-108-1
6	22.4	49.8	2576	3	US-09-479-264-1
7	22	48.9	22671	3	US-08-976-259-14
8	22	48.9	22671	3	US-09-956-004-14
9	21.8	48.4	10554	3	US-09-949-016-12320
10	21.8	48.4	10554	3	US-09-949-016-16084
11	21.8	48.4	51291	3	US-09-949-016-15362
12	21.6	48.0	1230230	3	US-09-198-452A-1
13	21.6	48.0	1230230	3	US-09-438-185A-1
14	21.4	47.6	2403	3	US-09-489-039A-20
15	21.4	47.6	4266	3	US-09-489-039A-7
16	21.2	47.1	601	3	US-09-949-016-72448
17	21.2	47.1	819	3	US-09-601-198-45
18	21.2	47.1	132871	3	US-09-949-016-13863
19	21.2	47.1	161607	3	US-09-949-016-12210
20	21.2	47.1	6706890	3	US-09-949-016-12505
21	21.2	47.1	6706890	3	US-09-949-016-14207
22	21	46.7	2749	3	US-09-619-353-11
23	21	46.7	2923	3	US-09-619-353-9
24	20.8	46.2	381	3	US-09-328-352-1745

C 25	20.8	46.2	966	2	US-08-181-271A-24	Sequence 24, Appl
C 26	20.8	46.2	966	2	US-08-449-315-24	Sequence 24, Appl
C 27	20.8	46.2	966	2	US-08-444-803-24	Sequence 24, Appl
C 28	20.8	46.2	966	2	US-08-449-043-24	Sequence 24, Appl
C 29	20.8	46.2	966	2	US-08-456-263A-24	Sequence 24, Appl
C 30	20.8	46.2	966	2	US-08-455-416-24	Sequence 24, Appl
C 31	20.8	46.2	966	2	US-08-455-244-24	Sequence 24, Appl
C 32	20.8	46.2	966	2	US-08-454-876-24	Sequence 24, Appl
C 33	20.8	46.2	966	2	US-08-457-364-24	Sequence 24, Appl
C 34	20.8	46.2	966	2	US-08-456-262-24	Sequence 24, Appl
C 35	20.8	46.2	966	2	US-08-456-240-24	Sequence 24, Appl
C 36	20.8	46.2	966	2	US-08-455-736-24	Sequence 24, Appl
C 37	20.8	46.2	966	2	US-08-971-217-24	Sequence 24, Appl
C 38	20.8	46.2	966	3	US-09-350-600-24	Sequence 24, Appl
C 39	20.8	46.2	966	3	US-09-906-234-24	Sequence 24, Appl
C 40	20.8	46.2	1204	2	US-08-181-271A-21	Sequence 21, Appl
C 41	20.8	46.2	1204	2	US-08-449-315-21	Sequence 21, Appl
C 42	20.8	46.2	1204	2	US-08-444-803-21	Sequence 21, Appl
C 43	20.8	46.2	1204	2	US-08-449-043-21	Sequence 21, Appl
C 44	20.8	46.2	1204	2	US-08-456-265A-21	Sequence 21, Appl
C 45	20.8	46.2	1204	2	US-08-455-416-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-134-000C-3345
Sequence 3345, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3345
LENGTH: 720
TYPR: DNA
ORGANISM: Enterococcus faecalis
US-09-134-000C-3345

Query Match 51.6%; Score 23.2; DB 3; Length 720;
Best Local Similarity 77.8%; Pred. No. 7;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGA 37
DB 36 GAAAGATTTCGCGATTCCTTCTTGATTAAGA 71

RESULT 2
US-09-949-016-156329/C
Sequence 156329, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASES, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 156329
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-156329

```

```

Query Match          51.1%; Score 23; DB 3; Length 601;
Best Local Similarity 74.4%; Pred. No. 8.1;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

QY      5 AGAATTCGGCCACCGCTTCTTGATTAAGAATGTTT 43
DB      554 ACAGTTCCCGCATCTCTTCTGATCAAGAGGTTT 516

```

```

RESULT 3
US-09-949-016-16108/c
; Sequence 16108, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16108
; LENGTH: 101300
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(101300)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16108

```

```

Query Match          51.1%; Score 23; DB 3; Length 101300;
Best Local Similarity 74.4%; Pred. No. 29;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

QY      5 AGAATTCGGCCACCGCTTCTTGATTAAGAATGTTT 43
DB      36639 ACAGTTCCCGCATCTCTTCTGATCAAGAGGTTT 36601

```

```

RESULT 4
US-09-513-999C-26046
; Sequence 26046, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; PATENT REFERENCE: 59,US2,REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 26046

```

```

; LENGTH: 364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 172
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 178
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 179
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 181
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 182
; OTHER INFORMATION: s=g or c
US-09-513-999C-26046

```

```

Query Match          50.2%; Score 22.6; DB 3; Length 364;
Best Local Similarity 75.7%; Pred. No. 10;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

QY      9 TTCCGCGCACCGCTTCTTGATTAAGAATGTTTTC 45
DB      88 TTCCCGCATCTCTTCTTGATCAAGAGGTTTTC 124

```

```

RESULT 5
US-09-265-108-1/c
; Sequence 1, Application US/09265108
; Patent No. 6033891
; GENERAL INFORMATION:
; APPLICANT: Gollightly, Elizabeth
; APPLICANT: Brown, Kimberly
; TITLE OF INVENTION: Nucleic Acids Encoding Polypeptides
; FILE REFERENCE: 5850.000-US
; CURRENT APPLICATION NUMBER: US/09/265,108
; PRIOR FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2576
; TYPE: DNA
; ORGANISM: Humicola
US-09-265-108-1

```

```

Query Match          49.8%; Score 22.4; DB 3; Length 2576;
Best Local Similarity 72.5%; Pred. No. 21;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

```

QY      1 CGAGAGATTCGGCCACCGCTTCTTGATTAAGAATG 40
DB      1555 CGAGAGATGACGCCACCGCCTCATGAGCTTAACAATG 1516

```

```

RESULT 6
US-09-479-264-1/c
; Sequence 1, Application US/09479264
; Patent No. 6280976
; GENERAL INFORMATION:
; APPLICANT: Elizabeth J. Gollightly
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Nucleic Acids Encoding Polypeptides
; FILE REFERENCE: 5850.200-US

```

;; CURRENT APPLICATION NUMBER: US/09/479,264
;; CURRENT FILING DATE: 2000-01-05
;; EARLIER APPLICATION NUMBER: 09/265,108
;; EARLIER FILING DATE: 1999-03-09
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1
;; LENGTH: 2576
;; TYPE: DNA
;; ORGANISM: Humicola
US-09-479-264-1

Query Match 49.8%; Score 22.4; DB 3; Length 2576;
Best Local Similarity 72.5%; Pred. No. 21;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CGACAGATTCCCGCCGCTTCTTGATTAAGATG 40
DB 1555 CGAAGATGACGCGCCCTCATTTGAGCTTAACATG 1516

RESULT 7

US-08-976-259-14/c
;; Sequence 14, Application US/08976259
;; Patent No. 6316609
;; GENERAL INFORMATION:
;; APPLICANT: Dillon, Patrick J.
;; APPLICANT: Choi, Gil H.
;; APPLICANT: Welch, Rodney A.
;; TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli
;; Patent No. 6316609
;; NUMBER OF SEQUENCES: 142
;; CORRESPONDENCE ADDRESS:
;; ADDRESSER: Sterne, Kessler, Goldstein & Fox P.L.L.C.
;; STREET: 1100 New York Ave, N.W., Suite 600
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20005-3934

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
;; COMPUTER: HP Vectra 486/33
;; OPERATING SYSTEM: MSDOS version 6.2
;; SOFTWARE: ASCII Text
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/976,259
;; FILING DATE: Herewith
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 60/031,626 AND US 60/061,953
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Steffe, Eric K.
;; REGISTRATION NUMBER: 36,688
;; REFERENCE/DOCKET NUMBER: 1488.0740002/EKS/CBM
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 371-2600
;; TELEFAX: (202) 371-2540
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 22671 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
US-08-976-259-14

Query Match 48.9%; Score 22; DB 3; Length 22671;
Best Local Similarity 73.7%; Pred. No. 52;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 8 ATTCCGACCGCTTCTTGATTAAGATGTTTC 45
DB 4970 ACTCCCAACGCTTCTTGCAAGTTTAAGTTTC 4933

RESULT 8
US-09-956-004-14/c
;; Sequence 14, Application US/09956004
;; Patent No. 6787643
;; GENERAL INFORMATION:
;; APPLICANT: Patrick J. Dillon et al.
;; TITLE OF INVENTION: Nucleotide Sequences of Escherichia coli Pathogenicity Islands
;; FILE REFERENCE: FB324DI
;; CURRENT APPLICATION NUMBER: US/09/956,004
;; CURRENT FILING DATE: 2001-09-20
;; PRIOR APPLICATION NUMBER: 08/976,259
;; PRIOR FILING DATE: 1997-11-21
;; PRIOR APPLICATION NUMBER: 60/061,953
;; PRIOR FILING DATE: 1997-10-14
;; PRIOR APPLICATION NUMBER: 60/031,626
;; PRIOR FILING DATE: 1996-11-22
;; NUMBER OF SEQ ID NOS: 142
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 14
;; LENGTH: 22671
;; TYPE: DNA
;; ORGANISM: Escherichia coli
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (19750)..(19750)
;; OTHER INFORMATION: n equals a, t, g, or c
;; NAME/KEY: misc feature
;; LOCATION: (20174)..(20174)
;; OTHER INFORMATION: n equals a, t, g, or c
US-09-956-004-14

Query Match 48.9%; Score 22; DB 3; Length 22671;
Best Local Similarity 73.7%; Pred. No. 52;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 8 ATTCCGACCGCTTCTTGATTAAGATGTTTC 45
DB 4970 ACTCCCAACGCTTCTTGCAAGTTTAAGTTTC 4933

RESULT 9
US-09-949-016-12320
;; Sequence 12320, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 12320
;; LENGTH: 10554
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-12320

Query Match 48.4%; Score 21.8; DB 3; Length 10554;
Best Local Similarity 78.8%; Pred. No. 52;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 AATTCGACCGCTTCTTGATTAAGAT 39
DB 1500 AATTCACCACTCTTATCTGTATTAATAAT 1532

```

RESULT 10
US-09-949-016-16084
; Sequence 16084, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16084
; LENGTH: 10554
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16084

```

```

Query Match      48.4%; Score 21.8; DB 3; Length 10554;
Best Local Similarity 78.8%; Pred. No. 52;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

Db      7 AATTCGACCGCTTCTGTGATTAAGAT 39
      1500 AATTCACCACTTATCTGTAATTAATAAT 1532

```

```

RESULT 11
US-09-949-016-15362
; Sequence 15362, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15362
; LENGTH: 51291
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)..(51291)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15362

```

```

Query Match      48.4%; Score 21.8; DB 3; Length 51291;
Best Local Similarity 70.7%; Pred. No. 77;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```

```

Db      2 GACAGATTCCGACCGCTTCTGTGATTAAGAT 42
      38731 GAAAGATTCAATTCACATTCATGATTAAGAACT 38771

```

```

RESULT 12
US-09-198-452A-1
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffee, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; thereof and uses thereof, in particular for the diagnosis, prevention
; and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; PRIOR FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t

```



```
; NAME/KEY: misc_feature
; LOCATION: (270001) ..(285000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (285001) ..(300000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (300001) ..(315000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (315001) ..(330000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (330001) ..(345000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (345001) ..(360000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (360001) ..(375000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (375001) ..(390000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (390001) ..(405000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (405001) ..(420000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (420001) ..(435000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (435001) ..(450000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (450001) ..(465000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (465001) ..(480000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (480001) ..(495000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (495001) ..(510000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (510001) ..(525000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (525001) ..(540000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (540001) ..(555000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (555001) ..(570000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (570001) ..(585000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (585001) ..(600000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (600001) ..(615000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (615001) ..(630000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
```

```
; LOCATION: (630001) ..(645000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (645001) ..(660000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (660001) ..(675000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (675001) ..(690000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (690001) ..(705000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (705001) ..(720000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (720001) ..(735000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (735001) ..(750000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (750001) ..(765000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (765001) ..(780000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (780001) ..(795000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (795001) ..(810000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (810001) ..(825000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (825001) ..(840000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (840001) ..(855000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (855001) ..(870000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (870001) ..(885000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (885001) ..(900000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (900001) ..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
```

```
Query Match 48.0%; Score 21.6; DB 3; Length 1230025;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 9 TTCCGCCACCGCTTTCCTTGATTAAGAAATGTTT 44
DB 520397 TTCTGCCATCTCTTGTGACAGACAAGATTGTTT 520432
```

```
RESULT 13
US-09-438-185A-1
; Sequence 1, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
```

APPLICANT: Mitchell, Wayne
APPLICANT: Kalman, Sue
APPLICANT: Davis, Ronald
TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
FILE REFERENCE: 018941-000411US
CURRENT APPLICATION NUMBER: US/09/438,185A
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: US 60/108,279
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: US 60/128,606
PRIOR FILING DATE: 1998-04-08
NUMBER OF SEQ ID NOS: 1074
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1230230
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
US-09-438-185A-1

Query Match 48.0%; Score 21.6; DB 3; Length 1230230;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9 TTCCGCCACCGCTTCTTGATTAAGATGTTT 44
DB 510309 TTCTGCCATCTCTTGCTAGAGCAAGATGTTT 510344

RESULT 14
US-09-489-039A-20
Sequence 20, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 20
LENGTH: 2403
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-20

Query Match 47.6%; Score 21.4; DB 3; Length 2403;
Best Local Similarity 71.8%; Pred. No. 52;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 AATTCGCCACCGCTTCTTGATTAAGATGTTTC 45
DB 2324 ACTTCGCTTCGCTTCGCTGATGATTCGATTTTTC 2362

RESULT 15
US-09-489-039A-7/c
Sequence 7, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7

LENGTH: 4266
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7

Query Match 47.6%; Score 21.4; DB 3; Length 4266;
Best Local Similarity 71.8%; Pred. No. 61;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 AATTCGCCACCGCTTCTTGATTAAGATGTTTC 45
DB 2024 ACTTCGCTTCGCTTCGCTGATGATTCGATTTTTC 1986

Search completed: April 7, 2006, 18:22:39
Job time : 21.0996 secs

C	1	45	100.0	45	3	US-09-211-691-4	Sequence 4, Appl
	2	45	100.0	45	6	US-10-317-773-4	Sequence 4, Appl
	3	45	100.0	45	6	US-10-317-428-4	Sequence 4, Appl
	4	27	60.0	623	9	US-10-915-740A-380	Sequence 380, Appl
C	5	27	60.0	65632	9	US-10-915-740A-49	Sequence 49, Appl
	6	27	60.0	2247716	9	US-10-915-740A-1068	Sequence 1068, Appl
	7	24	53.3	367	8	US-10-357-930-36950	Sequence 36950, A
	8	24	53.3	367	8	US-10-357-930-42575	Sequence 42575, A
C	9	23.6	52.4	31	8	US-10-819-861A-9	Sequence 9, Appl
	10	23.6	52.0	351	3	US-09-867-701-3041	Sequence 3041, Ap
C	11	22.8	50.7	3371	6	US-10-108-260A-1181	Sequence 1181, Ap
	12	22.4	49.8	1536	7	US-10-021-323-11510	Sequence 11510, A
	13	22.4	49.8	1035	4	US-09-925-065A-37059	Sequence 37059, A
	14	22.2	49.3	619	5	US-10-027-633-102109	Sequence 102109, A
	15	22.2	49.3	619	5	US-10-027-633-102110	Sequence 102110, A
	16	22.2	49.3	619	6	US-10-027-633-102109	Sequence 102109, A
	17	22.2	49.3	619	6	US-10-027-633-102110	Sequence 102110, A
C	18	22	48.9	551	4	US-09-925-065A-364339	Sequence 364339, A
	19	22	48.9	2914	7	US-10-437-963-83308	Sequence 83308, A
	20	22	48.9	6483	3	US-09-880-107-33699	Sequence 33699, Ap
	21	22	48.9	6483	3	US-10-435-696-24	Sequence 24, Appl
C	22	22	48.9	22671	7	US-09-956-604-14	Sequence 14, Appl
	23	22	48.9	22671	8	US-10-808-570-14	Sequence 14, Appl

APPLICANT: National Research Council of Canada

```
/ TITLE OF INVENTION: GalNac Transferase
/ FILE REFERENCE: 019633-000812US
/ CURRENT APPLICATION NUMBER: US/10/317,773
/ CURRENT FILING DATE: 2002-12-11
/ PRIOR APPLICATION NUMBER: US 60/069,443
/ PRIOR FILING DATE: 1997-12-15
/ PRIOR APPLICATION NUMBER: US 09/211,691
/ PRIOR FILING DATE: 1998-12-14
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4
/ LENGTH: 45
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:SYNM-R6 3'
US-10-317-773-4
```

```
Query Match          100.0%; Score 45; DB 6; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 CGACAGAAATTCGCGCACCGCTTTCCTTGATTAAGAAGTTTTC 45
Db      1 CGACAGAAATTCGCGCACCGCTTTCCTTGATTAAGAAGTTTTC 45
```

```
RESULT 3
US-10-317-428-4
/ Sequence 4, Application US/10317428
/ Publication No. US20030186414A1
/ GENERAL INFORMATION:
/ APPLICANT: Gilbert, Michel
/ APPLICANT: Young, N. Martin
/ APPLICANT: Wakarchuk, Warren W.
/ APPLICANT: National Research Council of Canada
/ TITLE OF INVENTION: Nucleic Acid That Encodes a Fusion Protein
/ FILE REFERENCE: 019633-000811US
/ CURRENT APPLICATION NUMBER: US/10/317,428
/ CURRENT FILING DATE: 2002-12-11
/ PRIOR APPLICATION NUMBER: US 60/069,443
/ PRIOR FILING DATE: 1997-12-15
/ PRIOR APPLICATION NUMBER: US 09/211,691
/ PRIOR FILING DATE: 1998-12-14
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4
/ LENGTH: 45
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:SYNM-R6 3'
US-10-317-428-4
```

```
Query Match          100.0%; Score 45; DB 6; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 CGACAGAAATTCGCGCACCGCTTTCCTTGATTAAGAAGTTTTC 45
Db      1 CGACAGAAATTCGCGCACCGCTTTCCTTGATTAAGAAGTTTTC 45
```

```
RESULT 4
US-10-915-740A-380
/ Sequence 380, Application US/10915740A
/ Publication No. US20050191316A1
/ GENERAL INFORMATION:
/ APPLICANT: Frazer, Claire M.
/ APPLICANT: Hickey, Erin
/ APPLICANT: Peterson, Jeremy
```

```
/ APPLICANT: Tetteijn, Herve
/ APPLICANT: Venter, J. Craig
/ APPLICANT: Massignani, Vega
/ APPLICANT: Galeotti, Cesira
/ APPLICANT: Mora, Manroza
/ APPLICANT: Ratti, Giulio
/ APPLICANT: Scarselli, Maria
/ APPLICANT: Scarlato, Vincenzo
/ APPLICANT: Rappuoli, Rino
/ APPLICANT: Pizza, Mariagrazia
/ APPLICANT: Grandi, Guido
/ TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
/ FILE REFERENCE: 002441.00090
/ CURRENT APPLICATION NUMBER: US/10/915,740A
/ CURRENT FILING DATE: 2004-08-11
/ PRIOR APPLICATION NUMBER: 09/806,866
/ PRIOR FILING DATE: 1999-10-08
/ PRIOR APPLICATION NUMBER: USSN 60/103,794
/ PRIOR FILING DATE: 1998-10-09
/ PRIOR APPLICATION NUMBER: USSN 60/132,068
/ PRIOR FILING DATE: 1999-04-30
/ PRIOR APPLICATION NUMBER: PCT/US99/25373
/ PRIOR FILING DATE: 1999-10-08
/ NUMBER OF SEQ ID NOS: 1068
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 380
/ LENGTH: 623
/ TYPE: DNA
/ ORGANISM: Neisseria meningitidis
US-10-915-740A-380
```

```
Query Match          60.0%; Score 27; DB 9; Length 623;
Best Local Similarity 100.0%; Pred. No. 0.84;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      19 GCTTCCCTTGATTAAGAAGTTTTC 45
Db      161 GCTTCCCTTGATTAAGAAGTTTTC 187
```

```
RESULT 5
US-10-915-740A-49/C
/ Sequence 49, Application US/10915740A
/ Publication No. US20050191316A1
/ GENERAL INFORMATION:
/ APPLICANT: Frazer, Claire M.
/ APPLICANT: Hickey, Erin
/ APPLICANT: Peterson, Jeremy
/ APPLICANT: Tetteijn, Herve
/ APPLICANT: Venter, J. Craig
/ APPLICANT: Massignani, Vega
/ APPLICANT: Galeotti, Cesira
/ APPLICANT: Mora, Manroza
/ APPLICANT: Ratti, Giulio
/ APPLICANT: Scarselli, Maria
/ APPLICANT: Scarlato, Vincenzo
/ APPLICANT: Rappuoli, Rino
/ APPLICANT: Pizza, Mariagrazia
/ APPLICANT: Grandi, Guido
/ TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
/ FILE REFERENCE: 002441.00090
/ CURRENT APPLICATION NUMBER: US/10/915,740A
/ CURRENT FILING DATE: 2004-08-11
/ PRIOR APPLICATION NUMBER: 09/806,866
/ PRIOR FILING DATE: 1999-10-08
/ PRIOR APPLICATION NUMBER: USSN 60/103,794
/ PRIOR FILING DATE: 1998-10-09
/ PRIOR APPLICATION NUMBER: USSN 60/132,068
/ PRIOR FILING DATE: 1999-04-30
/ PRIOR APPLICATION NUMBER: PCT/US99/25373
/ PRIOR FILING DATE: 1999-10-08
/ NUMBER OF SEQ ID NOS: 1068
/ SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 49
; LENGTH: 65632
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-915-740A-49

Query Match      60.0%; Score 27; DB 9; Length 65632;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 GCTTTCCTTGATTAAGATGTTTC 45
Db      55738 GCTTTCCTTGATTAAGATGTTTC 55712

RESULT 6
US-10-915-740A-1068
; Sequence 1068, Application US/10915740A
; Publication No. US20050191316A1
; GENERAL INFORMATION:
; APPLICANT: Frazer, Claire M.
; APPLICANT: Hickey, Erin
; APPLICANT: Peterson, Jeremy
; APPLICANT: Tetteilin, Herve
; APPLICANT: Venter, J. Craig
; APPLICANT: Massignani, Vega
; APPLICANT: Galeotti, Cesira
; APPLICANT: Mora, Marroa
; APPLICANT: Ratti, Giulio
; APPLICANT: Scarselli, Maria
; APPLICANT: Scariato, Vincenzo
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizzi, Mariagrazia
; APPLICANT: Grandi, Guido
; TITLE OF INVENTION: Neisseria Genomic Sequences And Methods Of Their Use
; FILE REFERENCE: 002441.00090
; CURRENT APPLICATION NUMBER: US/10/915,740A
; PRIOR FILING DATE: 2004-08-11
; PRIOR APPLICATION NUMBER: 09/806,866
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: USSN 60/103,794
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: USSN 60/132,068
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/US99/25373
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 1068
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1068
; LENGTH: 2242716
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-915-740A-1068

Query Match      60.0%; Score 27; DB 9; Length 2242716;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 GCTTTCCTTGATTAAGATGTTTC 45
Db      78420 GCTTTCCTTGATTAAGATGTTTC 78446

RESULT 7
US-10-357-930-36950
; Sequence 36950, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
```

```
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36950
; LENGTH: 367
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 340..341
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-36950

Query Match      53.3%; Score 24; DB 8; Length 367;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      4 CAGATTCGCCACCGCTTCTTGATTAAGATGTT 43
Db      271 CAGATTCGCCCTACTTCTTGATTAAGATGTT 310

RESULT 8
US-10-357-930-42575
; Sequence 42575, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42575
; LENGTH: 367
; TYPE: DNA
; ORGANISM: Homo sapiens
```

FEATURE:
NAME/KEY: misc feature
LOCATION: 340_341
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-42575

Query Match
Best Local Similarity 53.3%; Score 24; DB 8; Length 367;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CAGAAATCCGCCACCGCTTCTCTGTGATTAAAGATGTTT 43
Db 271 CACAAATTTAGCCTTACTTCTCTGTGATTAAAGAAATGTTT 310

RESULT 9
US-10-819-861A-9
Sequence 9, Application US/10819861A
Publication No. US20040241190A1
GENERAL INFORMATION:
APPLICANT: Xenova Research Limited
APPLICANT: McLean, Cornelia S.
APPLICANT: Keen, Simon W.
APPLICANT: Martin, Gillian May
TITLE OF INVENTION: Vaccine Preparations
FILE REFERENCE: 5673-68171-01
CURRENT APPLICATION NUMBER: US/10/819,861A
CURRENT FILING DATE: 2004-04-06
PRIOR APPLICATION NUMBER: GB 0308691.5
PRIOR FILING DATE: 2003-04-07
PRIOR APPLICATION NUMBER: US 60/464,758
PRIOR FILING DATE: 2003-04-21
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.2
SEQ ID NO 9
LENGTH: 31
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide primer.
US-10-819-861A-9

Query Match
Best Local Similarity 52.4%; Score 23.6; DB 8; Length 31;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 10 TCCGCCACCGCTTCTCTGTGATTAAAGAT 39
Db 2 TCGATATCGCTTCTCTGTGATTAAAGAT 31

RESULT 10
US-09-867-701-3041/C
Sequence 3041, Application US/09867701
Patent No. US2002013237A1
GENERAL INFORMATION:
APPLICANT: Aglate, Paul A.
APPLICANT: Jones, Robert
APPLICANT: Haillocker, Susan L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.497
CURRENT APPLICATION NUMBER: US/09/867,701
CURRENT FILING DATE: 2001-05-29
NUMBER OF SEQ ID NOS: 10912
SOFTWARE: FastSeq for windows Version 4.0
SEQ ID NO 3041
LENGTH: 351
TYPE: DNA
ORGANISM: Homo sapien
US-09-867-701-3041

Query Match
Best Local Similarity 52.0%; Score 23.4; DB 3; Length 351;

Best Local Similarity 73.2%; Pred. No. 23;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GACAGAAATCCGCCACCGCTTCTCTGTGATTAAAGATGTT 42
Db 240 GACAGAAATCCGCCACCGCTTCTCTGTGATTAAAGATGTT 200

RESULT 11
US-10-108-260A-1181
Sequence 1181, Application US/10108260A
Publication No. US20040005560A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. US20040005560A1el full length cDNA
FILE REFERENCE: H1-A0106
CURRENT APPLICATION NUMBER: US/10/108,260A
CURRENT FILING DATE: 2002-03-27
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1181
LENGTH: 3371
TYPE: DNA
ORGANISM: Homo sapiens
US-10-108-260A-1181

Query Match
Best Local Similarity 50.7%; Score 22.8; DB 6; Length 3371;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 3 ACAGAAATCCGCCACCGCTTCTCTGTGATTAAAGATGTTT 44
Db 1378 ACAGAAATCCGCCACCGCTTCTCTGTGATTAAAGATGTTT 1419

RESULT 12
US-10-021-323-11510
Sequence 11510, Application US/10021323
Publication No. US20040123340A1
GENERAL INFORMATION:
APPLICANT: Deikman, Jill
APPLICANT: Feng, Paul C.C.
APPLICANT: Fincher, Karen L.
APPLICANT: Ziegler, Todd E.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(52274)B
CURRENT APPLICATION NUMBER: US/10/021,323
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: US 60/255, 619
PRIOR FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 17880
SEQ ID NO 11510
LENGTH: 546
TYPE: DNA
ORGANISM: Gossypium hirsutum
FEATURE:
OTHER INFORMATION: Clone ID: LIB3829-002-01-K6-H1
US-10-021-323-11510

Query Match
Best Local Similarity 49.8%; Score 22.4; DB 7; Length 546;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 5 AGAATTCGCCACCGCTTCTCTGTGATTAAAGATGTTT 44
Db 80 AGAATTCGCCACCGCTTCTCTGTGATTAAAGATGTTT 119

RESULT 13
US-09-925-065A-37059
Sequence 37059, Application US/09925065A
Publication No. US20050228172A9

```
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 37059
/ LENGTH: 1035
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-925-065A-37059

Query Match
Best Local Similarity 49.8%; Score 22.4; DB 4; Length 1035;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 4 CAGAAATCCGCCACCGCTTCTTGATTAAGAATTT 43
DB 871 CCGAATTCCTCCAGAGTTCCCTTGATTAATTTATTT 910

RESULT 14
US-10-027-632-102109
/ Sequence 102109, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 102109
/ LENGTH: 619
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(619)
/ OTHER INFORMATION: n = A,T,C or G
US-10-027-632-102109

Query Match
Best Local Similarity 49.3%; Score 22.2; DB 5; Length 619;
Matches 27; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
```

```
Matches 27; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 3 ACAGAAATCCGCCACCGCTTCTTGATTAAGAAT 39
DB 114 AAAGGCTAACGCTCCACCTTCTTGATTAAGAAAT 150

RESULT 15
US-10-027-632-102110
/ Sequence 102110, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 102110
/ LENGTH: 619
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(619)
/ OTHER INFORMATION: n = A,T,C or G
US-10-027-632-102110

Query Match
Best Local Similarity 49.3%; Score 22.2; DB 5; Length 619;
Matches 27; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Search completed: April 7, 2006, 18:56:34
Job time : 76.7527 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 18:05:06 ; Search time 52.5328 Seconds
(without alignments)
3426.906 Million cell updates/sec

Title: US-09-211-691-4

Perfect score: 45
Sequence: 1 GACAGATTCGCCGCCGCGC.....tttgatgaagaatgttttc 45

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.New:*
1: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /SIDS5/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq1:*
7: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
8: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
9: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq3:*
10: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
11: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
12: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
13: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
14: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
15: /SIDS5/ptodata/1/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	23.4	52.0	159497	14 US-11-112-908-61	Sequence 61, Appl
C 2	23.4	52.0	171427	14 US-11-112-908-60	Sequence 60, Appl
C 3	22.4	49.8	1035	6 US-09-925-065A-37059	Sequence 37059, A
C 4	22.4	49.8	1035	9 US-10-301-480-138297	Sequence 138297, A
C 5	22.4	49.8	1035	10 US-10-301-480-751706	Sequence 751706, A
C 6	22.4	49.8	1035	10 US-10-301-480-434999	Sequence 434999, A
C 7	22.4	49.8	1035	10 US-10-301-480-1048408	Sequence 1048408, A
C 8	22.4	49.8	1035	6 US-09-925-065A-364339	Sequence 364339, A
C 9	21.8	48.4	200	14 US-11-098-686-7073	Sequence 7073, Ap
C 10	21.8	48.4	412	6 US-09-925-065A-134253	Sequence 134253, A
C 11	21.8	48.4	412	6 US-10-301-480-10543	Sequence 10543, A
C 12	21.8	48.4	412	10 US-10-301-480-623952	Sequence 623952, A
C 13	21.8	48.4	414	6 US-09-925-065A-179113	Sequence 179113, A
C 14	21.8	48.4	419	10 US-10-301-480-229997	Sequence 229997, A
C 15	21.8	48.4	419	10 US-10-301-480-843406	Sequence 843406, A
C 16	21.8	48.4	435	9 US-10-301-480-10542	Sequence 10542, A
C 17	21.8	48.4	435	10 US-10-301-480-623951	Sequence 623951, A
C 18	21.8	48.4	581	10 US-10-301-480-333515	Sequence 333515, A

19	21.8	48.4	581	10	US-10-301-480-946924	Sequence 946924, A
20	21.8	48.4	588	6	US-09-925-065A-254602	Sequence 254602, A
21	21.8	48.4	588	6	US-09-925-065A-254603	Sequence 254603, A
22	21.8	48.4	588	10	US-10-301-480-333516	Sequence 333516, A
23	21.8	48.4	588	10	US-10-301-480-946925	Sequence 946925, A
24	21.8	48.4	616	6	US-09-925-065A-854928	Sequence 854928, A
25	21.8	48.4	1000	10	US-10-301-480-580482	Sequence 580482, A
26	21.8	48.4	1000	10	US-10-301-480-1193891	Sequence 1193891, A
27	21.8	48.4	4167	14	US-11-098-686-9804	Sequence 9804, Ap
28	21.8	48.4	145619	14	US-11-098-686-8739	Sequence 8739, Ap
C 29	21.6	48.0	571	6	US-09-925-065A-195289	Sequence 195289, A
C 30	21.6	48.0	580	10	US-10-301-480-284500	Sequence 284500, A
C 31	21.6	48.0	580	10	US-10-301-480-997909	Sequence 997909, A
C 32	21.6	48.0	837	6	US-09-925-065A-578556	Sequence 578556, A
C 33	21.6	48.0	1380	9	US-10-301-480-25075	Sequence 25075, A
C 34	21.6	48.0	1380	10	US-10-301-480-838484	Sequence 838484, A
35	21.6	48.0	37415	9	US-10-330-773-486	Sequence 486, Ap
36	21.4	47.6	457	6	US-09-925-065A-443892	Sequence 443892, A
37	21.4	47.6	457	10	US-10-301-480-504230	Sequence 504230, A
38	21.4	47.6	457	10	US-10-301-480-1117639	Sequence 1117639, A
39	21.4	47.6	575	6	US-09-925-065A-391550	Sequence 391550, A
40	21.4	47.6	585	10	US-10-301-480-459770	Sequence 459770, A
41	21.4	47.6	585	10	US-10-301-480-1073179	Sequence 1073179, A
42	21.4	47.6	994	10	US-10-301-480-566796	Sequence 566796, A
43	21.4	47.6	994	10	US-10-301-480-1180205	Sequence 1180205, A
44	21.2	47.1	819	6	US-09-925-065A-4742	Sequence 4742, Ap
45	21.2	47.1	819	6	US-09-925-065A-4743	Sequence 4743, Ap

ALIGNMENTS

RESULT 1
US-11-112-908-61/c
; Sequence 61, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT FILING DATE: 2005-04-22
; PRIOR FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 61
; LENGTH: 159497
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-61
Query Match 52.0%; Score 23.4; DB 14; Length 159497;
Best Local Similarity 73.2%; Pred. No. 76;
Matches 30; Conservative 0; Mismatches 11; Indels 0;
Gaps 0;
C 2 GACGAGATTCGCCGCCGCTTCTTGATTAAGATGTT 42
DB 127197 GACGAGATTCGCCGCCGCTTCTTGATTAAGATGTT 127157
US-11-112-908-60/c
; Sequence 60, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:

```

; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 60
; LENGTH: 171427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-60

Query Match      52.0%; Score 23.4; DB 14; Length 171427;
Best Local Similarity 73.2%; Pred. No. 78;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

Qy      2 GACGAGATTCGGCCACCGCTTCTTGATTAAGAATGTT 42
Db      33333 GACGAGATTCGGCTCCACTGCTCTCTGATTAAGAAGCTT 33293
```

```

RESULT 3
US-09-925-065A-37059
; Sequence 37059, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37059
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-37059
```

```

Query Match      49.8%; Score 22.4; DB 6; Length 1035;
Best Local Similarity 72.5%; Pred. No. 46;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

Qy      4 CAGAAATTCGGCCACCGCTTCTTGATTAAGAATGTT 43
Db      871 CTGAATTCCTCCAGAGTTTCCTTGATTAATTCATTT 910
```

```

RESULT 4
US-10-301-480-138297
; Sequence 138297, Application US/10301480
; Publication No. US20060057564A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138297
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-138297

Query Match      49.8%; Score 22.4; DB 9; Length 1035;
Best Local Similarity 72.5%; Pred. No. 46;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

Qy      4 CAGAAATTCGGCCACCGCTTCTTGATTAAGAATGTT 43
Db      871 CTGAATTCCTCCAGAGTTTCCTTGATTAATTCATTT 910
```

```

RESULT 5
US-10-301-480-751706
; Sequence 751706, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 751706
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-751706
```

```

Query Match      49.8%; Score 22.4; DB 10; Length 1035;
Best Local Similarity 72.5%; Pred. No. 46;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

Qy      4 CAGAAATTCGGCCACCGCTTCTTGATTAAGAATGTT 43
Db      871 CTGAATTCCTCCAGAGTTTCCTTGATTAATTCATTT 910
```

```

RESULT 6
US-10-301-480-434999/C
; Sequence 434999, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
```

```
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 434999
/ LENGTH: 544
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-434999

Query Match
Best Local Similarity 48.9%; Score 22; DB 10; Length 544;
Pred. No. 55;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 8 ATTCGCGCACCGCTTCTTGTGATTAGAAATGTTTC 45
DB 508 ATTCCCTTGTCTTCTTGTGATTAGAAATGTTTC 471

RESULT 7
US-10-301-480-1048408/c
/ Sequence 1048408, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1048408
/ LENGTH: 544
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-1048408

Query Match
Best Local Similarity 48.9%; Score 22; DB 10; Length 544;
Pred. No. 55;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 8 ATTCGCGCACCGCTTCTTGTGATTAGAAATGTTTC 45
DB 508 ATTCCCTTGTCTTCTTGTGATTAGAAATGTTTC 471

RESULT 8
US-09-925-065A-364339/c
/ Sequence 364339, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
```

```
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 364339
/ LENGTH: 551
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-364339

Query Match
Best Local Similarity 48.9%; Score 22; DB 6; Length 551;
Pred. No. 55;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 8 ATTCGCGCACCGCTTCTTGTGATTAGAAATGTTTC 45
DB 508 ATTCCCTTGTCTTCTTGTGATTAGAAATGTTTC 471

RESULT 9
US-11-098-686-7073/c
/ Sequence 7073, Application US/11098686
/ Publication No. US20060024696A1
/ GENERAL INFORMATION:
/ APPLICANT: Kapur, Vivek and Gebhart, Connie J.
/ TITLE OF INVENTION: NUCLEIC ACID AND PEPTIDE SEQUENCES
/ FILE REFERENCE: 09531-128001
/ CURRENT APPLICATION NUMBER: US/11/098,686
/ PRIOR FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: PCT/US03/31318
/ PRIOR FILING DATE: 2003-10-01
/ PRIOR APPLICATION NUMBER: US 60/416,395
/ PRIOR FILING DATE: 2002-10-04
/ NUMBER OF SEQ ID NOS: 11433
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7073
/ LENGTH: 200
/ TYPE: DNA
/ ORGANISM: Lawsonia intracellularis
US-11-098-686-7073

Query Match
Best Local Similarity 48.4%; Score 21.8; DB 14; Length 200;
Pred. No. 50;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGAAATTCGCGCACCGCTTCTTGTGATTAGAAATGTTT 44
DB 151 CAGAAATTCGCGCACCGCTTCTTGTGATTAGAAATGTTT 111

RESULT 10
US-09-925-065A-134253
/ Sequence 134253, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 134253
/ LENGTH: 412
```

TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-134253

Query Match 48.4%; Score 21.8; DB 6; Length 412;
Best Local Similarity 70.7%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATGTT 42
DB 321 GCCAGATTCCAGCAAGCTTACTTTGATATGAAGGCT 361

RESULT 11
US-10-301-480-10543
Sequence 10543, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10543
LENGTH: 412
TYPE: DNA
ORGANISM: Homo sapiens
US-10-301-480-10543

Query Match 48.4%; Score 21.8; DB 9; Length 412;
Best Local Similarity 70.7%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATGTT 42
DB 321 GCCAGATTCCAGCAAGCTTACTTTGATATGAAGGCT 361

RESULT 12

US-10-301-480-623952
Sequence 623952, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 623952
LENGTH: 412
TYPE: DNA
ORGANISM: Homo sapiens
US-10-301-480-623952

Query Match 48.4%; Score 21.8; DB 10; Length 412;
Best Local Similarity 70.7%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATGTT 42

DB 321 GCCAGATTCCAGCAAGCTTACTTTGATATGAAGGCT 361

RESULT 13
US-09-925-065A-179113
Sequence 179113, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 179113
LENGTH: 414
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-179113

Query Match 48.4%; Score 21.8; DB 6; Length 414;
Best Local Similarity 70.7%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATGTT 42
DB 321 GCCAGATTCCAGCAAGCTTACTTTGATATGAAGGCT 361

RESULT 14

US-10-301-480-229997
Sequence 229997, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 229997
LENGTH: 419
TYPE: DNA
ORGANISM: Homo sapiens
US-10-301-480-229997

Query Match 48.4%; Score 21.8; DB 10; Length 419;
Best Local Similarity 70.7%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 GACAGATTCCGCCACCGCTTCTTGATTAAGATGTT 42
DB 321 GCCAGATTCCAGCAAGCTTACTTTGATATGAAGGCT 361

RESULT 15
 US-10-301-480-843406
 ; Sequence 843406, Application US/10301480
 ; Publication No. US20060057564A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
 ; FILE REFERENCE: 108827.137
 ; CURRENT APPLICATION NUMBER: US/10/301,480
 ; PRIOR FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 10/215,598
 ; PRIOR FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: US 60/311,695
 ; PRIOR FILING DATE: 2001-08-10
 ; NUMBER OF SEQ ID NOS: 1226818
 ; SOFTWARE: PatSeq for Windows Version 4.0
 ; SEQ ID NO 843406
 ; LENGTH: 419
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-10-301-480-843406

Query Match 48.4%; Score 21.8; DB 10; Length 419;
 Best Local Similarity 70.7%; Pred. No. 62;
 Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 2 GACGAGATTCGCGACCGCTTCTCTTGATTAAGATGTT 42
 Db 321 GCCGCGATTCGCGAAGCTTACTTTGATTAATGAAGGT 361

Search completed: April 7, 2006, 21:23:53
 Job time : 55.5328 secs

THIS PAGE BLANK (USPTO)